Figure 1

E H(B) E E E E EH H   E   E   H   B B S (B)

0.00   9.95 12.10   23.10   45.90

(B) Bg E P X E   E Bg E P K E H H Bg P K Bg K P X E P K P E   H

0   2.0   4.0   6.0   8.0   10.0 12.0 14.0 16.0 18.0 20.0 22.0 23.1   pPR1230

——— pPR1248

—1

——— pPR1245

— pPR1238

— pPR1240

—2

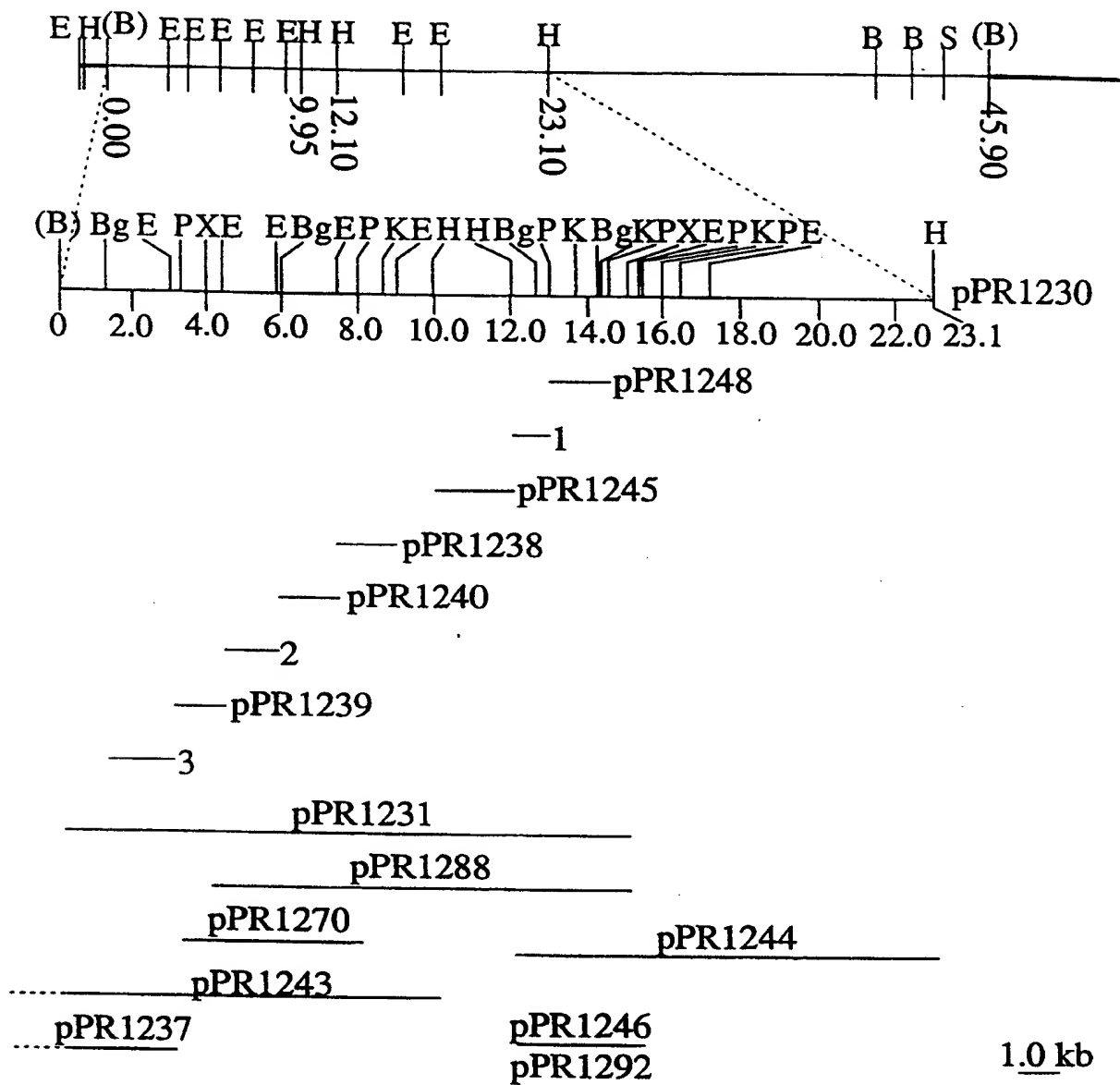—pPR1239

—3

pPR1231

pPR1288

pPR1270

pPR1244

pPR1243

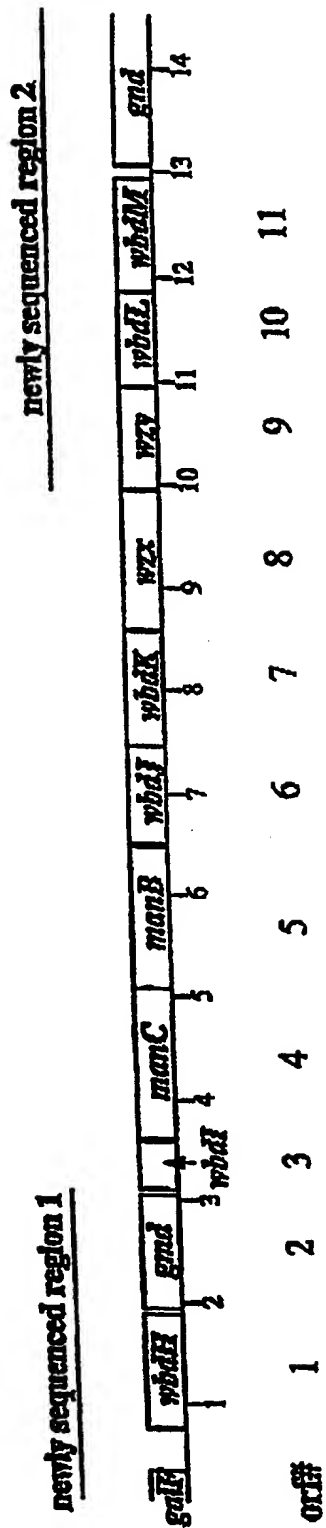pPR1237

pPR1246
pPR1292

1.0 kb

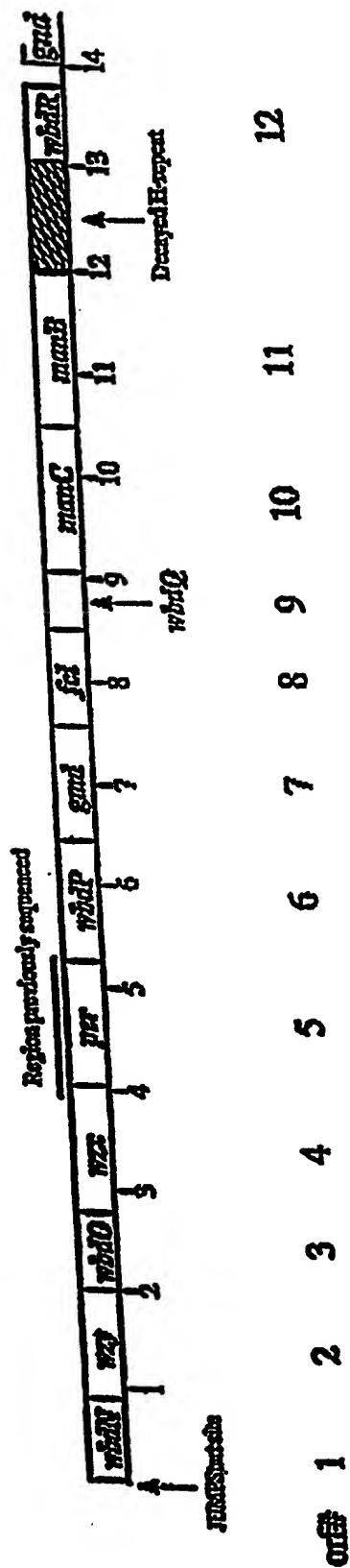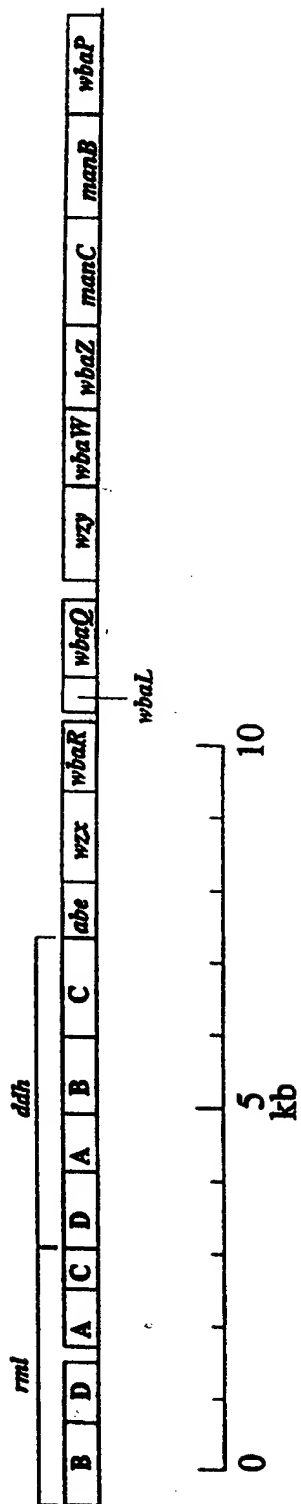# Figure 2

3/58

Figure 3
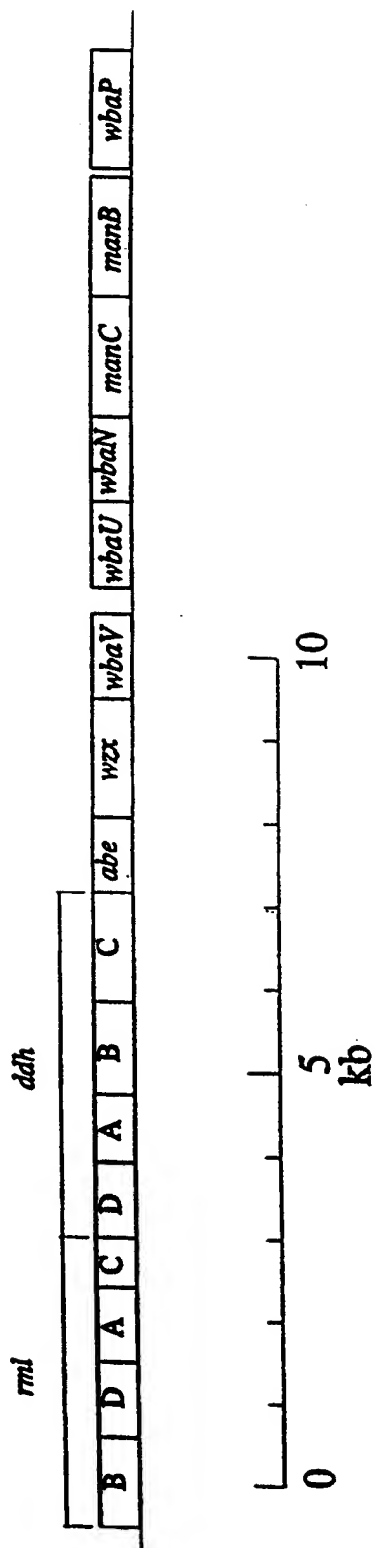
Figure 4

Figure 5

Figure 6

GATCTGATGGCCGTAGGGCGCTACGTGCTTTCTGCTGATATCTGGGCTGAGTTGGAAAAA   60

ACTGCTCCAGGTGCCTGGGGACGTATTCAACTGACTGATGCTATTGCAGAGTTGGCTAAA   120

AAACAGTCTGTTGATGCCATGCTGATGACCGGCGACAGCTACGACTGCGGTAAGAAGATG   180

GGCTATATGCAGGCATTCGTTAAGTATGGGCTGCGCAACCTTAAAGAAGGGGCGAAGTTC   240

CGTAAGAGCATCAAGAAGCTACTGAGTGAGTAGAGATTTACACGTCTTTGTGACGATAAG   300

CCAGAAAAAATAGCGGCAGTTAACATCCAGGCTTCTATGCTTTAAGCAATGGAATGTTAC   360

TGCCGTTTTTTATGAAAAATGACCAATAATAACAAGTTAACCTACCAAGTTTAATCTGCT   420

TTTTGTTGGATTTTTTCTTGTTTCTGGTCGCATTTGGTAAGACAATTAGCGTGAGTTTTA   480

GAGAGTTTTGCGGGATCTCGCGGAACTGCTCACATCTTTGGCATTTAGTTAGTGCACTGG   540

TAGCTGTTAAGCCAGGGGCGGTAGCTTGCCTAATTAATTTTTAACGTATACATTTATTCT   600

TGCCGCTTATAGCAAATAAAGTCAATCGGATTAAACTTCTTTTCCATTAGGTAAAAGAGT   660

GTTTGTAGTCGCTCAGGGAAATTGGTTTTGGTAGTAGTACTTTTCAAATTATCCATTTTC   720

Start of orf1
            M   L   L   C   C   I   H   I   N   V   Y   Y   L   L
CGATTTAGATGGCAGTTG<u>ATG</u>TTACTATGCTGCATACATATCAATGTATATTATTTACTT   780

   L   E   C   D   M   K   K   I   V   I   I   G   N   V   A   S   M   M   L   R
TTAGAATGTGATATGAAAAAAATAGTGATCATAGGCAATGTAGCGTCAATGATGTTAAGG   840

   F   R   K   E   L   I   M   N   L   V   R   Q   G   D   N   V   Y   C   L   A
TTCAGGAAAGAATTAATCATGAATTTAGTGAGGCAAGGTGATAATGTATATTGTCTAGCA   900

   N   D   F   S   T   E   D   L   K   V   L   S   S   W   G   V   K   G   V   K
AATGATTTTTCCACTGAAGATCTTAAAGTACTTTCGTCATGGGGCGTTAAGGGGGTTAAA   960

   F   S   L   N   S   K   G   I   N   P   F   K   D   I   I   A   V   Y   E   L
TTCTCTCTTAACTCAAAGGGTATTAATCCTTTTAAGGATATAATTGCTGTTTATGAACTA   1020

   K   K   I   L   K   D   I   S   P   D   I   V   F   S   Y   F   V   K   P   V
AAAAAAATTCTTAAGGATATTTCCCCAGATATTGTATTTTCATATTTTGTAAAGCCAGTA   1080

   I   F   G   T   I   A   S   K   L   S   K   V   P   R   I   V   G   M   I   E
ATATTTGGAACTATTGCTTCAAAGTTGTCAAAAGTGCCAAGGATTGTTGGAATGATTGAA   1140

   G   L   G   N   A   F   T   Y   Y   K   G   K   Q   T   T   K   T   K   M   I
GGTCTAGGTAATGCCTTCACTTATTATAAGGGAAAGCAGACCACAAAAACTAAAATGATA   1200

   K   W   I   Q   I   L   L   Y   K   L   A   L   P   M   L   D   D   L   I   L
AAGTGGATACAAATTCTTTTATATAAGTTAGCATTACCGATGCTTGATGATTTGATTCTA   1260

   L   N   H   D   D   K   K   D   L   I   D   Q   Y   N   I   K   A   K   V   T
TTAAATCATGATGATAAAAAAGATTTAATCGATCAGTATAATATTAAAGCTAAGGTAACA   1320

   V   L   G   G   I   G   L   D   L   N   E   F   S   Y   K   E   P   P   K   E
GTGTTAGGTGGGATTGGATTGGATCTTAATGAGTTTTCATATAAAGAGCCACCGAAAGAG   1380

   K   I   T   F   I   F   I   A   R   L   L   R   E   K   G   I   F   E   F   I
AAAATTACCTTTATTTTTATAGCAAGGTTATTAAGAGAGAAAGGGATATTTGAGTTTATT   1440

   E   A   A   K   F   V   K   T   T   Y   P   S   S   E   F   V   I   L   G   G
GAAGCCGCAAAGTTCGTTAAGACAACTTATCCAAGTTCTGAATTTGTAATTTTAGGAGGT   1500

Figure 7/1

```
        F  E  S  N  N  P  F  S  L  Q  K  N  E  I  E  S  L  R  K  E
        TTTGAGAGTAATAATCCTTTCTCATTACAAAAAAATGAAATTGAATCGCTAAGAAAAGAA      1560

        H  D  L  I  Y  P  G  H  V  E  N  V  Q  D  W  L  E  K  S  S
        CATGATCTTATTTATCCTGGTCATGTGGAAAATGTTCAAGATTGGTTAGAGAAAAGTTCT      1620

        V  F  V  L  P  T  S  Y  R  E  G  V  P  R  V  I  Q  E  A  M
        GTTTTTGTTTTACCTACATCATATCGAGAAGGCGTACCAAGGGTGATCCAAGAAGCTATG      1680

        A  I  G  R  P  V  I  T  T  N  V  P  G  C  R  D  I  I  N  D
        GCTATTGGTAGACCTGTAATAACAACTAATGTACCTGGGTGTAGGGATATAATAAATGAT      1740

        G  V  N  G  F  L  I  P  P  F  E  I  N  L  L  A  E  K  M  K
        GGGGTCAATGGCTTTTTGATACCTCCATTTGAAATTAATTTACTGGCAGAAAAAATGAAA      1800

        Y  F  I  E  N  K  D  K  V  L  E  M  G  L  A  G  R  K  F  A
        TATTTTATTGAGAATAAAGATAAAGTACTCGAAATGGGGCTTGCTGGAAGGAAGTTTGCA      1860

        E  K  N  F  D  A  F  E  K  N  N  R  L  A  S  I  I  K  S  N
        GAAAAAAACTTTGATGCTTTTGAAAAAAATAATAGACTAGCATCAATAATAAAATCAAAT      1920


            N  D  F  *     End of orf1
        AATGATTTT TGACTTGAGCAGAAATTATTTATATTTCAATCTGAAAAATAAAGGCTGTTA     1980

        Start of orf2
            M  N  K  V  A  L  I  T  G  I  T  G  Q  D  G  S  Y  L  A
        TTATGAATAAAGTGGCATTAATTACTGGTATCACTGGGCAAGATGGCTCCTATTTGGCAG     2040

        E  L  L  L  E  K  G  Y  E  V  H  G  I  K  R  R  A  S  S  F
        AATTATTGTTAGAAAAAGGTTATGAAGTTCATGGTATTAAACGCCGTGCATCTTCATTTA     2100

        N  T  E  R  V  D  H  I  Y  Q  D  S  H  L  A  N  P  K  L  F
        ATACTGAGCGAGTGGATCACATCTATCAGGATTCACATTTAGCTAATCCTAAACTTTTTC     2160

        L  H  Y  G  D  L  T  D  T  S  N  L  T  R  I  L  K  E  V  Q
        TACACTATGGCGATTTGACAGATACTTCCAATCTGACCCGTATTTTAAAAGAAGTTCAAC     2220

        P  D  E  V  Y  N  L  G  A  M  S  H  V  A  V  S  F  E  S  P
        CAGATGAAGTTTACAATTTGGGGGCGATGAGCCATGTAGCGGTATCATTTGAGTCACCAG     2280

        E  Y  T  A  D  V  D  A  I  G  T  L  R  L  L  E  A  I  R  I
        AATACACTGCTGATGTTGATGCGATAGGAACATTGCGTCTTCTTGAAGCTATCAGGATAT     2340

        L  G  L  E  K  K  T  K  F  Y  Q  A  S  T  S  E  L  Y  G  L
        TGGGGCTGGAAAAAAAGACAAAATTTTATCAGGCTTCAACTTCAGAGCTTTATGGTTTGG     2400

        V  Q  E  I  P  Q  K  E  T  T  P  F  Y  P  R  S  P  Y  A  V
        TTCAAGAAATTCCACAAAAAGAGACTACGCCATTTTATCCACGTTCGCCTTATGCTGTTG     2460

        A  K  L  Y  A  Y  W  I  T  V  N  Y  R  E  S  Y  G  M  F  A
        CAAAATTATATGCCTATTGGATCACTGTTAATTATCGTGAGTCTTATGGTATGTTTGCCT     2520

        C  N  G  I  L  F  N  H  E  S  P  R  R  G  E  T  F  V  T  R
        GCAATGGTATTCTCTTTAACCACGAATCACCTCGCCGTGGCGAGACCTTTGTTACTCGTA     2580

        K  I  T  R  G  I  A  N  I  A  Q  G  L  D  K  C  L  Y  L  G
        AAATAACACGCGGGATAGCAAATATTGCTCAAGGTCTTGATAAATGCTTATACTTGGGAA     2640

        N  M  D  S  L  R  D  W  G  H  A  K  D  Y  V  K  M  Q  W  M
        ATATGGATTCTCTGCGTGATTGGGGACATGCTAAGGATTATGTCAAAATGCAATGGATGA     2700
```

Figure 7/2

```
 M  L  Q  Q  E  T  P  E  D  F  V  I  A  T  G  I  Q  Y  S  V
TGCTGCAGCAAGAAACTCCAGAAGATTTTGTAATTGCTACAGGAATTCAATATTCTGTCC      2760

 R  E  F  V  T  M  A  A  E  Q  V  G  I  E  L  A  F  E  G  E
GTGAGTTTGTCACAATGGCGGCAGAGCAAGTAGGCATAGAGTTAGCATTTGAAGGTGAGG      2820

 G  V  N  E  K  G  V  V  V  S  V  N  G  T  D  A  K  A  V  N
GAGTAAATGAAAAAGGTGTTGTTGTTTCGGTCAATGGCACTGATGCTAAAGCTGTAAACC      2880

 P  G  D  V  I  I  S  V  D  P  R  Y  F  R  P  A  E  V  E  T
CGGGCGATGTAATTATATCTGTAGATCCAAGGTATTTTAGGCCTGCAGAAGTTGAAACCT      2940

 L  L  G  D  P  T  N  A  H  K  K  L  G  W  S  P  E  I  T  L
TGCTTGGCGATCCTACTAATGCGCATAAAAAATTAGGATGGAGCCCTGAAATTACATTGC      3000

 R  E  M  V  K  E  M  V  S  S  D  L  A  I  A  K  K  N  V  L
GTGAAATGGTAAAAGAAATGGTTTCCAGCGATTTAGCAATAGCGAAAAGAACGTCTTGC      3060

                                              End of orf2
 L  K  A  N  N  I  A  T  N  I  P  Q  E     *
TGAAAGCTAATAACATTGCCACTAATATTCCGCAAGAATAAAAAGATAATACATTAAAT ——3120


                                              Start of orf3
                                                  M  F
AATTAAAAATGGTGCTAGATTTATTAGTACCATTATTTTTTTTTGGGTGACTAATGTTTA      3180

 I  T  S  D  K  F  R  E  I  I  K  L  V  P  L  V  S  I  D  L
TTACATCAGATAAATTTAGAGAAATTATCAAGTTAGTTCCATTAGTATCAATTGATCTGC      3240

 L  I  E  N  E  N  G  E  Y  L  F  G  L  R  N  N  R  P  A  K
TAATTGAAAACGAGAATGGTGAATATTTATTTGGTCTTAGGAATAATCGACCGGCCAAAA      3300

 N  Y  F  F  V  P  G  G  R  I  R  K  N  E  S  I  K  N  A  F
ATTATTTTTTTGTTCCAGGTGGTAGGATTCGCAAAAATGAATCTATTAAAAATGCTTTTA      3360

 K  R  I  S  S  M  E  L  G  K  E  Y  G  I  S  G  S  V  F  N
AAAGAATATCATCTATGGAATTAGGTAAAGAGTATGGTATTTCAGGAAGTGTTTTTAATG      3420

 G  V  W  E  H  F  Y  D  D  G  F  F  S  E  G  E  A  T  H  Y
GTGTATGGGAACATTTCTATGATGATGGTTTTTTTTCTGAAGGCGAGGCAACACATTATA      3480

 I  V  L  C  Y  T  L  K  V  L  K  S  E  L  N  L  P  D  D  Q
TAGTGCTTTGTTACACACTGAAAGTTCTTAAAAGTGAATTGAATCTCCCAGATGATCAAC      3540

 H  R  E  Y  L  W  L  T  K  H  Q  I  N  A  K  Q  D  V  H  N
ATCGTGAATACCTTTGGCTAACTAAACACCAAATAAATGCTAAACAAGATGTTCATAACT      3600

               End of orf3                    Start of orf4
 Y  S  K  N  Y  F  L     *                              M
ATTCAAAAAATTATTTTTTGTAATTTTTATTAAAAATTAATATGCGAGAGAATTGTATGT      3660

 S  Q  C  L  Y  P  V  I  I  A  G  G  T  G  S  R  L  W  P  L
CTCAATGTCTTTACCCTGTAATTATTGCCGGAGGAACCGGAAGCCGTCTATGGCCGTTGT      3720

 S  R  V  L  Y  P  K  Q  F  L  N  L  V  G  D  S  T  M  L  Q
CTCGAGTATTATACCCTAAACAATTTTTAAATTTAGTTGGGGATTCTACAATGTTGCAA      3780

 T  T  I  T  R  L  D  G  I  E  C  E  N  P  I  V  I  C  N  E
CAACAATTACGCGTTTGGATGGCATCGAATGCGAAAATCCAATTGTTATCTGCAATGAAG      3840

 D  H  R  F  I  V  A  E  Q  L  R  Q  I  G  K  L  T  K  N  I
ATCACCGATTTATTGTAGCAGAGCAATTACGACAGATTGGTAAGCTAACCAAGAATATTA      3900

 I  L  E  P  K  G  R  N  T  A  P  A  I  A  L  A  A  F  I  A
TACTTGAGCCGAAAGGCCGTAATACTGCACCTGCCATAGCTTTAGCTGCTTTTATCGCTC      3960
```

Figure 7/3

WO 98/50531

PCT/AU98/00315

```
    Q   K   N   N   P   N   D   D   P   L   L   L   V   L   A   A   D   H   S   I
    AGAAGAATAATCCTAATGACGACCCTTTATTATTAGTACTTGCGGCAGACCACTCTATAA          4020

    N   N   E   K   A   F   R   E   S   I   I   K   A   M   P   Y   A   T   S   G
    ATAATGAAAAAGCATTTCGAGAGTCAATAATAAAAGCTATGCCGTATGCAACTTCTGGGA          4080

    K   L   V   T   F   G   I   I   P   D   T   A   N   T   G   Y   G   Y   I   K
    AGTTAGTAACATTTGGAATTATTCCGGACACGGCAAATACTGGTTATGGATATATTAAGA          4140

    R   S   S   S   A   D   P   N   K   E   F   P   A   Y   N   V   A   E   F   V
    GAAGTTCTTCAGCTGATCCTAATAAAGAATTCCCAGCATATAATGTTGCGGAGTTTGTAG          4200

    E   K   P   D   V   K   T   A   Q   E   Y   I   S   S   G   N   Y   Y   W   N
    AAAAACCAGATGTTAAAACAGCACAGGAATATATTTCGAGTGGGAATTATTACTGGAATA          4260

    S   G   M   F   L   F   R   A   S   K   Y   L   D   E   L   R   K   F   R   P
    GCCGGAATGTTTTTATTTCGCGCCAGTAAATATCTTGATGAACTACGGAAATTTAGACCAG          4320

    D   I   Y   H   S   C   E   C   A   T   A   T   A   N   I   D   M   D   F   V
    ATATTTATCATAGCTGTGAATGTGCAACCGCTACAGCAAATATAGATATGGACTTTGTCC          4380

    R   I   N   E   A   E   F   I   N   C   P   E   E   S   I   D   Y   A   V   M
    GAATTAACGAGGCTGAGTTTATTAATTGTCCTGAAGAGTCTATCGATTATGCTGTGATGG          4440

    E   K   T   K   D   A   V   V   L   P   I   D   I   G   W   N   D   V   G   S
    AAAAAACAAAAGACGCTGTAGTTCTTCCGATAGATATTGGCTGGAATGACGTGGGTTCTT          4500

    W   S   S   L   W   D   I   S   Q   K   D   C   H   G   N   V   C   H   G   D
    GGTCATCACTTTGGGATATAAGCCAAAAGGATTGCCATGGTAATGTGTGCCATGGGGATG          4560

    V   L   N   H   D   G   E   N   S   F   I   Y   S   E   S   S   L   V   A   T
    TGCTCAATCATGATGGAGAAAATAGTTTTATTTACTCTGAGTCAAGTCTGGTTGCGACAG          4620

    V   G   V   S   N   L   V   I   V   Q   T   K   D   A   V   L   V   A   D   R
    TCGGAGTAAGTAATTTAGTAATTGTCCAAACCAAGGATGCTGTACTGGTTGCGGACCGTG          4680

    D   K   V   Q   N   V   K   N   I   V   D   D   L   K   K   R   K   R   A   E
    ATAAAGTCCAAAATGTTAAAAACATAGTTGACGATCTAAAAAAGAGAAAACGTGCTGAAT          4740

    Y   Y   M   H   R   A   V   F   R   P   W   G   K   F   D   A   I   D   Q   G
    ACTACATGCATCGTGCAGTTTTTCGCCCTTGGGGTAAATTCGATGCAATAGACCAAGGCG          4800

    D   R   Y   R   V   K   K   I   I   V   K   P   G   E   G   L   D   L   R   M
    ATAGATATAGAGTAAAAAAAATAATAGTTAAACCAGGAGAAGGGTTAGATTTAAGGATGC          4860

    H   H   H   R   A   E   H   W   I   V   V   S   G   T   A   K   V   S   L   G
    ATCATCATAGGGCAGAGCATTGGATTGTTGTATCCGGTACTGCTAAAGTTTCACTAGGTA          4920

    S   E   V   K   L   L   V   S   N   E   S   I   Y   I   P   Q   G   A   K   Y
    GTGAAGTTAAACTATTAGTTTCTAATGAGTCTATATATATCCCTCAGGGAGCAAAATATA          4980

    S   L   E   N   P   G   V   I   P   L   H   L   I   E   V   S   S   G   D   Y
    GTCTTGAGAATCCAGGCGTAAATACCTTTGCATCTAATTGAAGTAAGTTCTGGTGATTACC          5040

    L   E   S   D   D   I   V   R   F   T   D   R   Y   N   S   K   Q   F   L   K
    TTGAATCAGATGATATAGTGCGTTTTACTGACAGATATAACAGTAAACAATTCCTAAAGC          5100
```

End of orf4 Start of orf5
```
                         M   N   K   I   T   C   F   K   A   Y   D   I   R   G   R   L
    R   D   *
    GAGATTGATAAATATGAATAAAATAACTTGCTTCAAAGCATATGATATACGTGGGCGTCT————5160
```

Figure 7/4

```
         G  A  E  L  N  D  E  I  A  Y  R  I  G  R  A  Y  G  E  F  F
        TGGTGCTGAATTGAATGATGAAATAGCATATAGAATTGGTCGCGCTTATGGTGAGTTTTT     5220

         K  P  Q  T  V  V  V  G  G  D  A  R  L  T  S  E  S  L  K  K
        TAAACCTCAAACTGTAGTTGTGGGAGGAGATGCTCGCTTAACAAGTGAGAGTTTAAAGAA     5280

          S  L  S  N  G  L  C  D  A  G  V  N  V  L  D  L  G  M  C  G
        ATCACTCTCAAATGGGCTATGTGATGCAGGCGTAAATGTCTTAGATCTTGGAATGTGTGG     5340

         T  E  E  I  Y  F  S  T  W  Y  L  G  I  D  G  G  I  E  V  T
        TACTGAAGAGATATATTTTTCCACTTGGTATTTAGGAATTGATGGTGGAATCGAGGTAAC     5400

         A  S  H  N  P  I  D  Y  N  G  M  K  L  V  T  K  G  A  R  P
        TGCAAGCCATAATCCAATTGATTATAATGGAATGAAATTAGTAACCAAAGGTGCTCGACC     5460

         I  S  S  D  T  G  L  K  D  I  Q  Q  L  V  E  S  N  N  F  E
        AATCAGCAGTGACACAGGTCTCAAAGATATACAACAATTAGTAGAGAGTAATAATTTTGA     5520

         E  L  N  L  E  K  K  G  N  I  T  K  Y  S  T  R  D  A  Y  I
        AGAGCTCAACCTAGAAAAAAAAGGGAATATTACCAAATATTCCACCCGAGATGCCTACAT     5580

         N  H  L  M  G  Y  A  N  L  Q  K  I  K  K  I  K  I  V  V  N
        AAATCATTTGATGGGCTATGCTAATCTGCAAAAAATAAAAAAAATCAAAATAGTTGTGAA     5640

         S  G  N  G  A  A  G  P  V  I  D  A  I  E  E  C  F  L  R  N
        TTCTGGGAATGGTGCAGCTGGTCCTGTTATTGATGCTATTGAGGAATGCTTTTTACGGAA     5700

         N  I  P  I  Q  F  V  K  I  N  N  T  P  D  G  N  F  P  H  G
        CAATATTCCGATTCAGTTTGTAAAAAATAAATAATACACCCGATGGTAATTTTCCACATGG     5760

         I  P  N  P  L  L  P  E  C  R  E  D  T  S  S  A  V  I  R  H
        TATCCCTAATCCATTACTACCTGAGTGCAGAGAAGATACCAGCAGTGCGGTTATAAGACA     5820

         S  A  D  F  G  I  A  F  D  G  D  F  D  R  C  F  F  F  D  E
        TAGTGCTGATTTTGGTATTGCATTTGATGGTGATTTTGATAGGTGTTTTTTCTTTGATGA     5880

         N  G  Q  F  I  E  G  Y  Y  I  V  G  L  L  A  E  V  F  L  G
        AAATGGACAATTTATTGAAGGATACTACATTGTTGGTTTATTAGCGGAAGTTTTTTTAGG     5940

         K  Y  P  N  A  K  I  I  H  D  P  R  L  I  W  N  T  I  D  I
        GAAATATCCAAACGCAAAAATCATTCATGATCCTCGCCTTATATGGAATACTATTGATAT     6000

         V  E  S  H  G  G  I  P  I  M  T  K  T  G  H  A  Y  I  K  Q
        CGTAGAAAGTCATGGTGGTATACCTATAATGACTAAAACCGGTCATGCTTACATTAAGCA     6060

         R  M  R  E  E  D  A  V  Y  G  G  E  M  S  A  H  H  Y  F  K
        AAGAATGCGTGAAGAGGATGCCGTATATGGCGGCGAAATGAGTGCGCATCATTATTTTAA     6120

         D  F  A  Y  C  D  S  G  M  I  P  W  I  L  I  C  E  L  L  S
        AGATTTTGCATACTGCGATAGTGGAATGATTCCTTGGATTTTAATTTGTGAACTTTTGAG     6180

         L  T  N  K  K  L  G  E  L  V  C  G  C  I  N  D  W  P  A  S
        TCTGACAAATAAAAAATTAGGTGAACTGGTTTGTGGTTGTATAAACGACTGGCCGGCAAG     6240

         G  E  I  N  C  T  L  D  N  P  Q  N  E  I  D  K  L  F  N  R
        TGGAGAAATAAACTGTACACTAGACAATCCGCAAAATGAAATAGATAAATTATTTAATCG     6300

         Y  K  D  S  A  L  A  V  D  Y  T  D  G  L  T  M  E  F  S  D
        TTACAAAGATAGTGCCTTAGCTGTTGATTACACTGATGGATTAACTATGGAGTTCTCTGA     6360

         W  R  F  N  V  R  C  S  N  T  E  P  V  V  R  L  N  V  E  S
        TTGGCGTTTTAATGTTAGATGCTCAAATACAGAACCTGTAGTACGATTGAATGTAGAATC     6420

         R  N  N  A  I  L  M  Q  E  K  T  E  E  I  L  N  F  I  S  K
        TAGGAATAATGCTATTCTTATGCAGGAAAAAACAGAAGAAATTCTGAATTTTATATCAAA     6480
```

Figure 7/5

End of orf5   Start of orf6

```
                                                      M  K  V  L  L  T  G
ATAAATTTGCACCTGAGTTCATAATGGGAACAAGAAATATATGAAAGTACTTCTGACTGG            6540

  S  T  G  M  V  G  K  N  I  L  E  H  D  S  A  S  K  Y  N  I
CTCAACTGGCATGGTTGGTAAGAATATATTAGAGCATGATAGTGCAAGTAAATATAATAT            6600

  L  T  P  T  S  S  D  L  N  L  L  D  K  N  E  I  E  K  F  M
ACTTACTCCAACCAGCTCTGATTTGAATTTATTAGATAAAAATGAAATAGAAAAATTCAT            6660

  L  I  N  M  P  D  C  I  I  H  A  A  G  L  V  G  G  I  H  A
GCTTATCAACATGCCAGACTGTATTATACATGCAGCGGGATTAGTTGGAGGCATTCATGC            6720

  N  I  S  R  P  F  D  F  L  E  K  N  L  Q  M  G  L  N  L  V
AAATATAAGCAGGCCGTTTGATTTTCTGGAAAAAAATTTGCAGATGGGTTTAAATTTAGT            6780

  S  V  A  K  K  L  G  I  K  K  V  L  N  L  G  S  S  C  M  Y
TTCCGTCGCAAAAAAACTAGGTATCAAGAAAGTGCTTAACTTGGGTAGTTCATGCATGTA            6840

  P  K  N  F  E  E  A  I  P  E  K  A  L  L  T  G  E  L  E  E
CCCCAAAAACTTTGAAGAGGCTATTCCTGAGAAAGCTCTGTTAACTGGTGAGCTAGAAGA            6900

  T  N  E  G  Y  A  I  A  K  I  A  V  A  K  A  C  E  Y  I  S
AACTAATGAGGGATATGCTATTGCGAAAATTGCTGTAGCAAAAGCATGCGAATATATATC            6960

  R  E  N  S  N  Y  F  Y  K  T  I  I  P  C  N  L  Y  G  K  Y
AAGAGAAAACTCTAATTATTTTTATAAAACAATTATCCCATGTAATTTATATGGGAAATA            7020

  D  K  F  D  D  N  S  S  H  M  I  P  A  V  I  K  K  I  H  H
TGATAAATTTGATGATAACTCGTCACATATGATTCCGGCAGTTATAAAAAAAATCCATCA            7080

  A  K  I  N  N  V  P  E  I  E  I  W  G  D  G  N  S  R  R  E
TGCCGAAAATTAATAATGTCCCAGAGATCGAAATTTGGGGGGATGGTAATTCGCGCCGTGA            7140

  F  M  Y  A  E  D  L  A  D  L  I  F  Y  V  I  P  K  I  E  F
GTTTATGTATGCAGAAGATTTAGCTGATCTTATTTTTTATGTTATTCCTAAAATAGAATT            7200

  M  P  N  M  V  N  A  G  L  G  Y  D  Y  S  I  N  D  Y  Y  K
CATGCCTAATATGGTAAATGCTGGTTTAGGTTACGATTATTCAATTAATGACTATTATAA            7260

  I  I  A  E  E  I  G  Y  T  G  S  F  S  H  D  L  T  K  P  T
GATAATTGCAGAAGAAATTGGTTATACTGGGAGTTTTTCTCATGATTTAACAAAACCAAC            7320

  G  M  K  R  K  L  V  D  I  S  L  L  N  K  I  G  W  S  S  H
AGGAATGAAACGGAAGCTAGTAGATATTTCATTGCTTAATAAAATTGGTTGGTCAAGTCA            7380

  F  E  L  R  D  G  I  R  K  T  Y  N  Y  Y  L  E  N  Q  N  K
CTTTGAACTCAGAGATGGCATCAGAAAGACCTATAATTATTACTTGGAGAATCAAAATAA            7440
```

Start of orf7, End of orf6

```
  M  I  T  Y  P  L  A  S  N  T  W  D  E  Y  E  Y  A  A  I  Q
  *
ATGATTACATACCCACTTGCTAGTAATACTTGGGATGAATATGAGTATGCAGCAATACAG            7500

  S  V  I  D  S  K  M  F  T  M  G  K  K  V  E  L  Y  E  K  N
TCAGTAATTGACTCAAAAAATGTTTACCATGGGTAAAAAGGTTGAGTTATATGAGAAAAT            7560

  F  A  D  L  F  G  S  K  Y  A  V  M  V  S  S  G  S  T  A  N
TTTGCTGATTTGTTTGGTAGCAAATATGCCGTAATGGTTAGCTCTGGTTCTACAGCTAAT            7620
```

Figure 7/6

```
  L  L  M  I  A  A  L  F  F  T  N  K  P  K  L  K  R  G  D  E
CTGTTAATGATTGCTGCCCTTTTCTTCACTAATAAACCAAAACTTAAAAGAGGTGATGAA      7680

  I  I  V  P  A  V  S  W  S  T  T  Y  Y  P  L  Q  Q  Y  G  L
ATAATAGTACCTGCAGTGTCATGGTCTACGACATATTACCCTCTGCAACAGTATGGCTTA      7740

  K  V  K  F  V  D  I  N  K  E  T  L  N  I  D  I  D  S  L  K
AAGGTGAAGTTTGTCGATATCAATAAAGAAACTTTAAATATTGATATCGATAGTTTGAAA      7800

  N  A  I  S  D  K  T  K  A  I  L  T  V  N  L  L  G  N  P  N
AATGCTATTTCAGATAAAACAAAAGCAATATTGACAGTAAATTTATTAGGTAATCCTAAT      7860

  D  F  A  K  I  N  E  I  I  N  N  R  D  I  I  L  L  E  D  N
GATTTTGCAAAAATAAATGAGATAATAAATAATAGGGATATTATCTTACTAGAAGATAAC      7920

  C  E  S  M  G  A  V  F  Q  N  K  Q  A  G  T  F  G  V  M  G
TGTGAGTCGATGGGCGCGGTCTTTCAAAATAAGCAGGCAGGCACATTCGGAGTTATGGGT      7980

  T  F  S  S  F  Y  S  H  H  I  A  T  M  E  G  G  C  V  V  T
ACCTTTAGTTCTTTTTACTCTCATCATATAGCTACAATGGAAGGGGGCTGCGTAGTTACT      8040

  D  D  E  E  L  Y  H  V  L  L  C  L  R  A  H  G  W  T  R  N
GATGATGAAGAGCTGTATCATGTATTGTTGTGCCTTCGAGCTCATGGTTGGACAAGAAAT      8100

  L  P  K  E  N  M  V  T  G  T  K  S  D  D  I  F  E  E  S  F
TTACCAAAAGAGAATATGGTTACAGGCACTAAGAGTGATGATATTTTCGAAGAGTCGTTT      8160

  K  F  V  L  P  G  Y  N  V  R  P  L  E  M  S  G  A  I  G  I
AAGTTTGTTTTACCAGGATACAATGTTCGCCCACTTGAAATGAGTGGTGCTATTGGGATA      8220

  E  Q  L  K  K  L  P  G  F  I  S  T  R  R  S  N  A  Q  Y  F
GAGCAACTTAAAAAGTTACCAGGTTTTATATCCACCAGACGTTCCAATGCACAATATTTT      8280

  V  D  K  F  K  D  H  P  F  L  D  I  Q  K  E  V  G  E  S  S
GTAGATAAATTTAAAGATCATCCATTCCTTGATATACAAAAAGAAGTTGGTGAAAGTAGC      8340

  W  F  G  F  S  F  V  I  K  E  G  A  A  I  E  R  K  S  L  V
TGGTTTGGTTTTTCCTTCGTTATAAAGGAGGGAGCTGCTATTGAGAGGAAGAGTTTAGTA      8400

  N  N  L  I  S  A  G  I  E  C  R  P  I  V  T  G  N  F  L  K
AATAATCTGATCTCAGCAGGCATTGAATGCCGACCAATTGTTACTGGGAATTTTCTCAAA      8460

  N  E  R  V  L  S  Y  F  D  Y  S  V  H  D  T  V  A  N  A  E
AATGAACGTGTTTTGAGTTATTTTGATTACTCTGTACATGATACGGTAGCAAATGCCGAA      8520

  Y  I  D  K  N  G  F  F  V  G  N  H  Q  I  P  L  F  N  E  I
TATATAGATAAGAATGGTTTTTTTGTCGGAAACCACCAGATACCTTTGTTTAATGAAATA      8580

  D  Y  L  R  K  V  L  K  *        End of orf7
GATTATCTACGAAAAGTATTAAAATAACTAACGAGGCACTCTATTTCGAATAGAGTGCCT      8640

       Start of orf8
        M  V  L  T  V  K  K  I  L  A  F  G  Y  S  K  V  L  P
TTAAGATGGTATTAACAGTGAAAAAAATTTTAGCGTTTGGCTATTCTAAAGTACTACCAC      8700

  P  V  I  E  Q  F  V  N  P  I  C  I  F  I  I  T  P  L  I  L
CGGTTATTGAACAGTTTGTCAATCCAATTTGCATCTTCATTATCACACCACTAATACTCA      8760

  N  H  L  G  K  Q  S  Y  G  N  W  I  L  L  I  T  I  V  S  F
ACCACCTGGGTAAGCAAAGCTATGGTAATTGGATTTTATTAATTACTATTGTATCTTTTT      8820
```

Figure 7/7

```
        S  Q  L  I  C  G  G  C  S  A  W  I  A  K  I  I  A  E  Q  R
        CTCAGTTAATATGTGGAGGATGTTCCGCATGGATTGCAAAAATCATTGCAGAACAGAGAA         8880

        I  L  S  D  L  S  K  K  N  A  L  R  Q  I  S  Y  N  F  S  I
        TTCTTAGTGATTTATCAAAAAAAAATGCTTTACGTCAAATTTCCTATAATTTTTCAATTG         8940

        V  I  I  A  F  A  V  L  I  S  F  L  I  L  S  I  C  F  F  D
        TTATTATCGCATTTGCGGTATTGATTTCTTTTCTTATATTAAGTATTTGTTCTTCGATG         9000

        V  A  R  N  N  S  S  F  L  F  A  I  I  I  C  G  F  F  Q  E
        TTGCGAGGAATAATTCTTCATTCTTATTCGCGATTATTATTTGTGGTTTTTTTCAGGAAG         9060

        V  D  N  L  F  S  G  A  L  K  G  F  E  K  F  N  V  S  C  F
        TTGATAATTTATTTAGTGGTGCGCTAAAAGGTTTTGAAAAATTTAATGTATCATGTTTTT         9120

        F  E  V  I  T  R  V  L  W  A  S  I  V  I  Y  G  I  Y  G  N
        TTGAAGTAATTACAAGAGTGCTCTGGGCTTCTATAGTAATATATGGCATTTACGGAAATG         9180

        A  L  L  Y  F  T  C  L  A  F  T  I  K  G  M  L  K  Y  I  L
        CACTCTTATATTTTACATGTTTAGCCTTTACCATTAAAGGTATGCTAAAATATATTCTTG         9240

        V  C  L  N  I  T  G  C  F  I  N  P  N  F  N  R  V  G  I  V
        TATGTCTGAATATTACCGGTTGTTTCATCAATCCTAATTTTAATAGAGTTGGGATTGTTA         9300

        N  L  L  N  E  S  K  W  M  F  L  Q  L  T  G  G  V  S  L  S
        ATTTGTTAAATGAGTCAAAATGGATGTTTCTTCAATTAACTGGTGGCGTCTCACTTAGTT         9360

        L  F  D  R  L  V  I  P  L  I  L  S  V  S  K  L  A  S  Y  V
        TGTTTGATAGGCTCGTAATACCATTGATTTTATCTGTCAGTAAACTGGCTTCTTATGTCC         9420

        P  C  L  Q  L  A  Q  L  M  F  T  L  S  A  S  A  N  Q  I  L
        CTTGCCTTCAACTAGCTCAATTGATGTTCACTCTTTCTGCGTCTGCAAATCAAATATTAC         9480

        L  P  M  F  A  R  M  K  A  S  N  T  F  P  S  N  C  F  F  K
        TACCAATGTTTGCTAGAATGAAAGCATCTAACACATTTCCCTCTAATTGTTTTTTTAAAA         9540

        I  L  L  V  S  L  I  S  V  L  P  C  L  A  L  F  F  F  G  R
        TTCTGCTTGTATCACTAATTTCTGTTTTGCCTTGTCTTGCGTTATTCTTTTTTGGTCGTG         9600

        D  I  L  S  I  W  I  N  P  T  F  A  T  E  N  Y  K  L  M  Q
        ATATATTATCAATATGGATAAACCCTACATTTGCAACTGAAAATTATAAATTAATGCAAA         9660

        I  L  A  I  S  Y  I  L  L  S  M  M  T  S  F  H  F  L  L  L
        TTTTAGCTATAAGTTACATTTTATTGTCAATGATGACATCTTTTCATTTCTTGTTATTAG         9720

        G  I  G  K  S  K  L  V  A  N  L  N  L  V  A  G  L  A  L  A
        GAATTGGTAAATCTAAGCTTGTTGCAAATTTAAATCTGGTTGCAGGGCTCGCACTTGCTG         9780

        A  S  T  L  I  A  A  H  Y  G  L  Y  A  I  S  M  V  K  I  I
        CTTCAACGTTAATCGCAGCTCATTATGGCCTTTATGCAATATCTATGGTAAAAATAATAT         9840

        Y  P  A  F  Q  F  Y  Y  L  Y  V  A  F  V  Y  F  N  R  A  K
        ATCCGGCTTTTCAATTTTATTACCTTTATGTAGCTTTTGTCTATTTTAATAGAGCGAAAA         9900
```

Start of orf9, End of orf8

```
          M  S  I  D  L  L  F  S  I  T  E  I  A  I  V  F  S  C  T  I
          N  V  Y  *
          ATGTCTATTGATTTACTTTTTTTCAATTACTGAAATCGCAATTGTTTTTTCTTGCACTATT         9960

            Y  I  F  T  Q  C  L  L  M  R  R  I  Y  L  D  K  S  I  L  I
            TACATATTTACTCAATGTTTGTTAATGCGGAGGATCTATTTAGATAAAAGTATTTTAATT       10020

            L  L  C  L  L  F  F  L  V  I  I  Q  L  P  E  L  N  V  N  G
            CTTTTATGCTTGCTCTTTTTTTTAGTAATCATTCAACTTCCTGAGCTTAATGTAAACGGT       10080
```

Figure 7/8

```
      L   V   D   S   L   K   L   S   L   P   L   L   M   V   F   I   A   F   Q   K
      TTGGTCGATTCTTTAAAGTTATCACTGCCTTTATTGATGGTCTTTATCGCTTTTCAAAAA        10140

      P   K   L   C   L   W   V   I   I   A   L   L   F   L   N   S   A   F   N   F
      CCGAAATTATGCTTGTGGGTTATTATTGCATTGTTGTTTTTGAACTCTGCATTTAATTTT        10200

      L   Y   L   K   T   F   D   K   F   S   S   F   P   F   T   F   F   I   L   L
      TTATATTTAAAGACATTCGATAAGTTTAGCTCATTTCCTTTTACTTTTTTTATATTGCTG        10260

      F   Y   L   F   R   L   G   I   G   N   L   P   V   Y   K   N   K   K   F   Y
      TTTTACTTGTTTAGATTGGGAATTGGTAATTTACCGGTTTATAAAAATAAAAAATTTTAC        10320

      A   L   I   F   L   F   I   L   I   D   I   M   Q   S   L   L   I   N   Y   R
      GCGTTGATTTTTCTCTTTATATTAATAGACATAATGCAGTCATTGTTAATAAATTATAGG        10380

      G   Q   I   L   Y   S   V   I   C   I   L   I   L   V   F   K   V   N   L   R
      GGGCAGATTTTATATTCCGTAATTTGCATCCTGATACTTGTGTTTAAAGTTAATTTAAGA        10440

      K   K   I   P   Y   F   F   L   M   L   P   V   L   Y   V   I   I   M   A   Y
      AAAAAGATTCCATACTTTTTTTTTAATGCTGCCAGTTTTATATGTAATTATTATGGCTTAT        10500

      I   G   F   N   Y   F   N   K   G   V   T   F   F   E   P   T   A   S   N   I
      ATTGGTTTTAATTATTTCAATAAAGGCGTAACTTTTTTTGAACCTACAGCAAGTAATATT        10560

      E   R   T   G   M   I   Y   Y   L   V   S   Q   L   G   D   Y   I   F   H   G
      GAACGTACGGGGATGATATATTATTTGGTTTCACAGCTTGGTGATTATATATTCCATGGT        10620

      M   G   T   L   N   F   L   N   N   G   G   Q   Y   K   T   L   Y   G   L   P
      ATGGGGACATTAAATTTCTTAAATAACGGCGGACAATATAAGACGTTATATGGACTTCCA        10680

      S   L   I   P   N   D   P   H   D   F   L   L   R   F   F   I   S   I   G   V
      TCATTAATTCCTAATGACCCTCATGATTTTTTATTACGGTTCTTTATAAGTATTGGTGTG        10740

      I   G   A   L   V   Y   H   S   I   F   F   V   F   F   R   R   I   S   F   L
      ATAGGAGCATTGGTTTATCATTCTATATTTTTTGTTTTTTTTAGGAGAATATCTTTCTTA        10800

      L   Y   E   R   N   A   P   F   I   V   V   S   C   L   L   L   L   Q   V   V
      TTATATGAGAGAAATGCTCCTTTCATTGTTGTAAGTTGTTTGTTACTGTTACAAGTTGTG        10860

      L   I   Y   T   L   N   P   F   D   A   F   N   R   L   I   C   G   L   T   V
      TTAATTTATACATTAAACCCTTTTGATGCTTTTAATCGATTGATTTGCGGGCTTACAGTT        10920
```

Start of orf10   **End of orf9**
```
      G   V   V   Y   G   F   A   K   I   R   *
                          M   D   L   Q   K   L   D   K   Y   T   C   N   G   N   L   D   A
      GGAGTTGTTTATGGATTTGCAAAAATTAGATAAGTATACCTGTAATGGAAATTTAGACGC        10980

      P   L   V   S   I   I   I   A   T   Y   N   S   E   L   D   I   A   K   C   L
      TCCACTTGTTTCAATAATCATTGCAACTTATAATTCTGAACTTGATATAGCTAAGTGTTT        11040

      Q   S   V   T   N   Q   S   Y   K   N   I   E   I   I   I   M   D   G   G   S
      GCAATCGGTAACTAATCAATCTTATAAGAATATTGAAATCATAATAATGGATGGAGGATC        11100

      S   D   K   T   L   D   I   A   K   S   F   K   D   D   R   I   K   I   V   S
      TTCTGATAAAACGCTTGATATTGCAAAATCGTTTAAAGACGACCGAATAAAAATAGTTTC        11160

      E   K   D   R   G   I   Y   D   A   W   N   K   A   V   D   L   S   I   G   D
      AGAGAAAGATCGTGGAATTTATGATGCCTGGAATAAAGCAGTTGATTTATCCATTGGTGA        11220

      W   V   A   F   I   G   S   D   D   V   Y   Y   H   T   D   A   I   A   S   L
      TTGGGTAGCATTTATTGGTTCAGATGATGTTTACTATCATACAGATGCAATTGCTTCATT        11280

      M   K   G   V   M   V   S   N   G   A   P   V   V   Y   G   R   T   A   H   E
      GATGAAGGGGGTTATGGTATCTAATGGCGCCCCTGTGGTTTATGGGAGGACAGCGCACGA        11340
```

Figure 7/9

```
     G   P   D   R   N   I   S   G   F   S   G   S   E   W   Y   N   L   T   G   F
   AGGTCCCGATAGGAACATATCTGGATTTTCAGGCAGTGAATGGTACAACCTAACAGGATT      11400

     K   F   N   Y   Y   K   C   N   L   P   L   P   I   M   S   A   I   Y   S   R
   TAAGTTTAATTATTACAAATGTAATTTACCATTGCCCATTATGAGCGCAATATATTCTCG      11460

     D   F   F   R   N   E   R   F   D   I   K   L   K   I   V   A   D   A   D   W
   TGATTTCTTCAGAAACGAACGTTTTGATATTAAATTAAAAATTGTTGCTGACGCTGATTG      11520

     F   L   R   C   F   I   K   W   S   E   K   S   P   Y   F   I   N   D   T
   GTTTCTGAGATGTTTCATCAAATGGAGTAAAGAGAAGTCACCTTATTTTATTAATGACAC      11580

     T   P   I   V   R   M   G   Y   G   G   V   S   T   D   I   S   S   Q   V   K
   GACCCCTATTGTTAGAATGGGATATGGTGGGGTTTCGACTGATATTTCTTCTCAAGTTAA      11640

     T   T   L   E   S   F   I   V   R   K   K   N   N   I   S   C   L   N   I   Q
   AACTACGCTAGAAAGTTTCATTGTACGCAAAAAGAATAATATATCCTGTTTAAACATACA      11700

     L   I   L   R   Y   A   K   I   L   V   M   V   A   I   K   N   I   F   G   N
   GCTGATTCTTAGATATGCTAAAATTCTGGTGATGGTAGCGATCAAAAATATTTTTGGCAA      11760

     N   V   Y   K   L   M   H   N   G   Y   H   S   L   K   K   I   K   N   K   I
   TAATGTTTATAAATTAATGCATAACGGGTATCATTCCCTAAAGAAAATCAAGAATAAAAT      11820
```

Start of orf11, End of orf10
```
     M   K   I   V   Y   I   I   T   G   L   T   C   G   G   A   E   H   L   M   T
     *
   ATGAAGATTGTTTATATAATAACCGGGCTTACTTGTGGTGGAGCCGAACACCTTATGACG      11880

     Q   L   A   D   Q   M   F   I   R   G   H   D   V   N   I   I   C   L   T   G
   CAGTTAGCAGACCAAATGTTTATACGCGGGCATGATGTTAATATTATTTGTCTAACTGGT      11940

     I   S   E   V   K   P   T   Q   N   I   N   I   H   Y   V   N   M   D   K   N
   ATATCTGAGGTAAAGCCAACACAAAATATTAATATTCATTATGTTAATATGGATAAAAAT      12000

     F   R   S   F   F   R   A   L   F   Q   V   K   K   I   I   V   A   L   K   P
   TTTAGAAGCTTTTTTAGAGCTTTATTTCAAGTAAAAAAAATAATTGTCGCCTTAAAGCCA      12060

     D   I   I   H   S   H   M   F   H   A   N   I   F   S   R   F   I   R   M   L
   GATATAATACATAGTCATATGTTTCATGCTAATATTTTTAGTCGTTTTATTAGGATGCTG      12120

     I   P   A   V   P   L   I   C   T   A   H   N   K   N   E   G   G   N   A   R
   ATTCCAGCGGTGCCCCTGATATGTACCGCACACAACAAAAATGAAGGTGGCAATGCAAGG      12180

     M   F   C   Y   R   L   S   D   F   L   A   S   I   T   T   N   V   S   K   E
   ATGTTTTGTTATCGACTGAGTGATTTTTTAGCTTCTATTACTACAAATGTAAGTAAAGAG      12240

     A   V   Q   E   F   I   A   R   K   A   T   P   K   N   K   I   V   E   I   P
   GCTGTTCAAGAGTTTATAGCAAGAAAGGCTACACCTAAAAATAAAATAGTAGAGATTCCG      12300

     N   F   I   N   T   N   K   F   D   F   D   I   N   V   R   K   K   T   R   D
   AATTTTATTAATACAAATAAATTTGATTTTGATATTAATGTCAGAAAGAAAACGCGAGAT      12360

     A   F   N   L   K   D   S   T   A   V   L   L   A   V   G   R   L   V   E   A
   GCTTTTAATTTGAAAGACAGTACAGCAGTACTGCTCGCAGTAGGAAGACTTGTTGAAGCA      12420

     K   D   Y   P   N   L   L   N   A   I   N   H   L   I   L   S   K   T   S   N
   AAAGACTATCCGAACTTATTAAATGCAATAAATCATTTGATTCTTTCAAAAACATCAAAT      12480

     C   N   D   F   I   L   L   I   A   G   D   G   A   L   R   N   K   L   L   D
   TGTAATGATTTTATTTTGCTTATTGCTGGCGATGGCGCATTAAGAAATAAATTATTGGAT      12540

     L   V   C   Q   L   N   L   V   D   K   V   F   F   L   G   Q   R   S   D   I
   TTGGTTTGTCAATTGAATCTTGTGGATAAAGTTTTCTTCTTGGGGCAAAGAAGTGATATT      12600
```

Figure 7/10

```
 K   E   L   M   C   A   A   D   L   F   V   L   S   S   E   W   E   G   F   G
AAAGAATTAATGTGTGCTGCAGATCTTTTTGTTTTGAGTTCTGAGTGGGAAGGTTTTGGT        12660

 L   V   V   A   E   A   M   A   C   E   R   P   V   V   A   T   D   S   G   G
CTCGTTGTTGCAGAAGCTATGGCGTGTGAACGTCCCGTTGTTGCTACCGATTCTGGTGGA        12720

 V   K   E   V   V   G   P   H   N   D   V   I   P   V   S   N   H   I   L   L
GTTAAAGAAGTCGTTGGACCTCATAATGATGTTATCCCTGTCAGTAATCATATTCTGTTG        12780

 A   E   K   I   A   E   T   L   K   I   D   D   N   A   R   K   I   I   G   M
GCAGAGAAAATCGCTGAGACACTTAAAATAGATGATAACGCAAGAAAAATAATAGGTATG        12840

 K   N   R   E   Y   I   V   S   N   F   S   I   K   T   I   V   S   E   W   E
AAAAATAGAGAATATATTGTTTCCAATTTTTCAATTAAAACGATAGTGAGTGAGTGGGAG        12900
```

```
                                                End of orf11
 R   L   Y   F   K   Y   S   K   R   N   N   I   I   D       *
CGCTTATATTTTAAATATTCCAAGCGTAATAATATAATTGAT TGAAAATATAAGTTTGTA        12960

CTCTGGATGCAATAGTTTCTCTATGCTGTTTTTTTACTGGCTCCGTATTTTTACTTATAG        13020

CTGGATTTTGTTATATATCAGTATTAATCTGTCTCAACTTCATCTAGACTACATTCAAGC        13080
```

```
                                              Start of gnd
                                               M   S   K   Q   Q   I
CGCGCATGCGTCGCGCGGTGACTACACCTGACAGGAGTATGTA ATGTCCAAGCAACAGAT        13140

 G   V   V   G   M   A   V   M   G   R   N   L   A   L   N   I   E   S   R   G
CGGCGTCGTCGGTATGGCAGTGATGGGGCGCAACCTGGCGCTCAACATCGAAAGCCGCGG        13200

 Y   T   V   S   I   F   N   R   S   R   E   K   T   E   E   V   V   A   E   N
TTATACCGTCTCCATCTTCAACCGCTCCCGCGAGAAAACTGAAGAAGTTGTTGCCGAGAA        13260

 P   D   K   K   L   V   P   Y   Y   T   V   K   E   F   V   E   S   L   E   T
CCCGGATAAGAAACTGGTTCCTTATTACACGGTGAAAGAGTTCGTCGAGTCTCTTGAAAC        13320

 P   R   R   I   L   L   M   V   K   A   G   A   G   T   D   A   A   I   D   S
CCCACGTCGTATCCTGTTAATGGTAAAAGCAGGGGCGGGAACTGATGCTGCTATCGATTC        13380

 L   K   P   Y   L   D   K   G   D   I   I   I   D   G   G   N   T   F   F   Q
CCTGAAGCCGTATCTGGATAAAGGCGACATCATTATTGATGGTGGCAACACCTTCTTCCA        13440

 D   T   I   R   R   N   R   E   L   S   A   E   G   F   N   F   I   G   T   G
GGACACTATCCGTCGTAACCGTGAACTGTCCGCGGAAGGCTTTAACTTCATCGGTACCGG        13500

 V   S   G   G   E   E   G   A   L   K   G   P   S   I   M   P   G   G   Q   K
CGTGTCCGGCGGTGAAGAGGGCGCCCTGAAAGGCCCATCTATCATGCCAGGTGGCCAGAA        13560

 E   A   Y   E   L   V   A   P   I   L   T   K   I   A   A   V   A   E   D   G
AGAAGCGTATGAGCTGGTTGCGCCTATCCTGACCAAGATTGCTGCGGTTGCTGAAGATGG        13620

 E   P   C   I   T   Y   I   G   A   D   G   A   G   H   Y   V   K   M   V   H
CGAACCATGTATAACTTACATCGGTGCTGACGGTGCGGGTCACTACGTGAAGATGGTGCA        13680

 N   G   I   E   Y   G   D   M   Q   L   I   A   E   A   Y   S   L   L   K   G
CAACGGTATCGAATATGGCGATATGCAGCTGATTGCTGAAGCCTATTCTCTGCTTAAAGG        13740

 G   L   N   L   S   N   E   E   L   A   T   T   F   T   E   W   N   E   G   E
CGGCCTTAATCTGTCTAACGAAGAGCTGGCAACCACTTTTACCGAGTGGAATGAAGGCGA        13800

 L   S   S   Y   L   I   D   I   T   K   D   I   F   T   K   K   D   E   E   G
GCTAAGTAGCTACCTGATTGACATCACCAAAGACATCTTCACCAAAAAAGATGAAGAGGG        13860
```

Figure 7/11

09/423093

```
 K  Y  L  V  D  V  I  L  D  E  A  A  N  K  G  T  G  K  W  T
TAAATACCTGGTTGATGTGATCCTGGACGAAGCTGCGAACAAAGGCACCGGTAAATGGAC      13920

 S  Q  S  S  L  D  L  G  E  P  L  S  L  I  T  E  S  V  F  A
CAGCCAGAGCTCTCTGGATCTGGGTGAACCGCTGTCGCTGATCACCGAATCCGTATTCGC      13980

 R  Y  I  S  S  L  K  D  Q  R  I  A  A  S  K  V  L  S  G  P
TCGCTACATCTCTTCTCTGAAAGACCAGCGCATTGCGGCATCTAAAGTGCTGTCTGGTCC      14040

 Q  A  K  L  A  G  D  K  A  E  F  V  E  K  V  R  R  A  L  Y
GCAGGCTAAACTGGCTGGTGATAAAGCAGAGTTCGTTGAGAAAGTCCGTCGCGCGCTGTA      14100

 L  G  K  I  V  S  Y  A  Q  G  F  S  Q  L  R  A  A  S  D  E
CCTGGGTAAAATCGTCTCTTATGCCCAAGGCTTCTCTCAACTGCGTGCCGCGTCTGACGA      14160

 Y  N  W  D  L  N  Y  G  E  I  A  K  I  F  R  A  G  C  I  I
ATACAACTGGGATCTGAACTACGGCGAAATCGCGAAGATCTTCCGCGCGGGCTGCATCAT      14220

 R  A  Q  F  L  Q  K  I  T  D  A  Y  A  E  N  K  G  I  A  N
TCGTGCGCAGTTCCTGCAGAAAATTACTGACGCGTATGCTGAAAACAAAGGCATTGCTAA      14280

 L  L  L  A  P  Y  F  K  N  I  A  D  E  Y  Q  Q  A  L  R  D
CCTGTTGCTGGCTCCGTACTTCAAAAATATCGCTGATGAATATCAGCAAGCGCTGCGTGA      14340

 V  V  A  Y  A  V  Q  N  G  I  P  V  P  T  F  S  A  A  V  A
TGTAGTGGCTTATGCTGTGCAGAACGGTATTCCGGTACCGACCTTCTCTGCAGCGGTAGC      14400

 Y  Y  D  S  Y  R  S  A  V  L  P  A  N  L  I  Q  A  Q  R  D
CTACTACGACAGCTACCGTTCTGCGGTACTGCCGGCTAATCTGATTCAGGCACAGCGTGA      14460

 Y  F  G  A  H  T  Y  K  R  T  D  K  E  G  V  F  H  T  G
TTACTTCGGTGCGCACACGTATAAACGCACTGATAAAGAAGGTGTGTTCCACACCG        14516
```

Figure 7/12

```
                 GTAACCAAGGGCGGTACGTGCATAAATTTTAATGCTTATCAAAACTATTAGCATTAAAAA      60


                      Start of orf1
                      M  N  K  E  T  V  S  I  I  M  P  V  Y  N
                 TATATAAGAAATTCTCAAATGAACAAAGAAACCGTTTCAATAATTATGCCCGTTTACAAT     120

                 G  A  K  T  I  I  S  S  V  E  S  I  I  H  Q  S  Y  Q  D  F
                 GGGGCCAAAACTATAATCTCATCAGTAGAATCAATTATACATCAATCTTATCAAGATTTT     180

                 V  L  Y  I  I  D  D  C  S  T  D  D  T  F  S  L  I  N  S  R
                 GTTTTGTATATCATTGACGATTGTAGCACCGATGATACATTTTCATTAATCAACAGTCGA     240

                 Y  K  N  N  Q  K  I  R  I  L  R  N  K  T  N  L  G  V  A  E
                 TACAAAAACAATCAGAAAATAAGAATATTGCGTAACAAGACAAATTTAGGTGTTGCAGAA     300

                 S  R  N  Y  G  I  E  M  A  T  G  K  Y  I  S  F  C  D  A  D
                 AGTCGAAATTATGGAATAGAAATGGCCACGGGGAAATATATTTCTTTTTGTGATGCGGAT     360

                 D  L  W  H  E  K  K  L  E  R  Q  I  E  V  L  N  N  E  C  V
                 GATTTGTGGCACGAGAAAAAATTAGAGCGTCAAATCGAAGTGTTAAATAATGAATGTGTA     420

                 D  V  V  C  S  N  Y  Y  V  I  D  N  N  R  N  I  V  G  E  V
                 GATGTGGTATGTTCTAATTATTATGTTATAGATAACAATAGAAATATTGTTGGCGAAGTT     480

                 N  A  P  H  V  I  N  Y  R  K  M  L  M  K  N  Y  I  G  N  L
                 AATGCTCCTCATGTGATAAATTATAGAAAAATGCTCATGAAAAACTACATAGGGAATTTG     540

                 T  G  I  Y  N  A  N  K  L  G  K  F  Y  Q  K  K  I  G  H  E
                 ACAGGAATCTATAATGCCAACAAATTGGGTAAGTTTTATCAAAAAAAGATTGGTCACGAG     600

                 D  Y  L  M  W  L  E  I  I  N  K  T  N  G  A  I  C  I  Q  D
                 GATTATTTGATGTGGCTGGAAATAATTAATAAAACAAATGGTGCTATTTGTATTCAAGAT     660

                 N  L  A  Y  Y  M  R  S  N  N  S  L  S  G  N  K  I  K  A  A
                 AATCTGGCGTATTACATGCGTTCAAATAATTCACTATCGGGTAATAAAATTAAAGCTGCA     720

                 K  W  T  W  S  I  Y  R  E  H  L  H  L  S  F  P  K  T  L  Y
                 AAATGGACATGGAGTATATATAGAGAACATTTACATTTGTCCTTTCCAAAAACATTATAT     780

                 Y  F  L  L  Y  A  S  N  G  V  M  K  K  I  T  H  S  L  L  R
                 TATTTTTTATTATATGCTTCAAATGGAGTCATGAAAAAAATAACACATTCACTATTAAGG     840


                      Start of orf2, End of orf1
                 R  K  E  T  K  K  *
                                   V  K  S  A  A  K  L  I  F  L  F  L  F  T
                 AGAAAGGAGACTAAAAAGTGAAGTCAGCGGCTAAGTTGATTTTTTTATTCCTATTTACAC     900

                 L  Y  S  L  Q  L  Y  G  V  I  I  D  D  R  I  T  N  F  D  T
                 TTTATAGTCTCCAGTTGTATGGGGTTATCATAGATGATCGTATAACAAATTTTGATACAA     960

                 K  V  L  T  S  I  I  I  I  F  Q  I  F  F  V  L  L  F  Y  L
                 AGGTATTAACTAGTATTATAATTATATTTCAGATTTTTTTTGTTTTATTATTTTATCTAA    1020

                 T  I  I  N  E  R  K  Q  Q  K  K  F  I  V  N  W  E  L  K  L
                 CGATTATAAATGAAAGAAAACAGCAGAAAAAATTTATCGTGAACTGGGAGCTAAAGTTAA    1080

                 I  L  V  F  L  F  V  T  I  E  I  A  A  V  V  L  F  L  K  E
                 TACTCGTTTTCCTTTTTGTGACTATAGAAATTGCTGCTGTAGTTTTATTTCTTAAAGAAG    1140

                 G  I  P  I  F  D  D  D  P  G  G  A  K  L  R  I  A  E  G  N
                 GTATTCCTATATTTGATGATGATCCAGGGGGGGCTAAACTTAGAATAGCTGAAGGTAATG    1200
```

Figure 8/1

```
      G   L   Y   I   R   Y   I   K   Y   F   G   N   I   V   V   F   A   L   I   I
     GACTTTACATTAGATATATTAAGTATTTTGGTAATATAGTTGTGTTTGCATTAATTATTC        1260

      L   Y   D   E   H   K   F   K   Q   R   T   I   I   F   V   Y   F   T   T   I
     TTTATGATGAGCATAAATTCAAACAGAGGACCATCATATTTGTATATTTTACAACGATTG        1320

      A   L   F   G   Y   R   S   E   L   V   L   L   I   L   Q   Y   I   L   I   T
     CTTTATTTGGTTATCGTTCTGAATTGGTGTTGCTCATTCTTCAATATATATTGATTACCA        1380

      N   I   L   S   K   D   N   R   N   P   K   I   K   R   I   I   G   Y   F   L
     ATATCCTGTCAAAGGATAACCGTAATCCTAAAATAAAAAGAATAATAGGGTATTTTTTAT        1440

      L   V   G   V   V   C   S   L   F   Y   L   S   L   G   Q   D   G   E   Q   N
     TGGTAGGGGTTGTATGCTCGTTGTTTTATCTAAGTTTAGGACAAGACGGAGAACAAAATG        1500

      D   S   Y   N   N   M   L   R   I   I   N   R   L   T   I   E   Q   V   E   G
     ACTCATATAATAATATGTTAAGGATAATTAATAGGTTAACAATAGAGCAAGTTGAAGGTG        1560

      V   P   Y   V   V   S   E   S   I   K   N   D   F   F   P   T   P   E   L   E
     TTCCATATGTTGTTTCTGAATCTATTAAGAACGATTTCTTTCCGACACCAGAGTTAGAAA        1620

      K   E   L   K   A   I   I   N   R   I   Q   G   I   K   H   Q   D   L   F   Y
     AGGAATTAAAAGCAATAATAAATAGAATACAGGGAATAAAGCATCAAGACTTATTTTATG        1680

      G   E   R   L   H   K   Q   V   F   G   D   M   G   A   N   F   L   S   V   T
     GAGAACGGTTACATAAACAAGTATTTGGAGACATGGGAGCAAATTTTTTATCAGTTACTA        1740

      T   Y   G   A   E   L   L   V   F   F   G   F   L   C   V   F   I   I   P   L
     CGTATGGAGCAGAACTGTTAGTTTTTTTTGGTTTTCTCTGTGTATTCATTATCCCTTTAG        1800

      G   I   Y   I   P   F   Y   L   L   K   R   M   K   K   T   H   S   S   I   N
     GGATATATATACCTTTTTATCTTTTAAAGAGAATGAAAAAAACCCATAGCTCGATAAATT        1860

      C   A   F   Y   S   Y   I   I   M   I   L   L   Q   Y   L   V   A   G   N   A
     GCGCATTCTATTCATATATCATTATGATTTTATTGCAATACTTAGTGGCTGGGAATGCAT        1920

      S   A   F   F   F   G   P   F   L   S   V   L   I   M   C   T   P   L   I   L
     CGGCCTTCTTTTTTGGTCCTTTTCTCTCCGTATTGATAATGTGTACTCCTCTGATCTTAT        1980


                                    Start of orf3
                                        M   K   I   S   V   I   T   V   T   Y
      L   H   D   T   L   K   R   L   S   R   N   E   N   I   S   Y   N   C   D   L
     TGCATGATACGTTAAAGAGATTATCACGAAATGAAAATATCAGTTATAACTGTGACTTAT        2040

   End of orf2
      N   N   A   E   G   L   E   K   T   L   S   S   L   S   I   L   K   I   K   P
    *
     AATAATGCTGAAGGGTTAGAAAAAACTTTAAGTAGTTTATCAATTTTAAAAATAAAACCT        2100

      F   E   I   I   I   V   D   G   G   S   T   D   G   T   N   R   V   I   S   R
     TTTGAGATTATTATAGTTGATGGCGGCTCTACAGATGGAACGAATCGTGTCATTAGTAGA        2160

      F   T   S   M   N   I   T   H   V   Y   E   K   D   E   G   I   Y   D   A   M
     TTTACTAGTATGAATATTACACATGTTTATGAAAAAGATGAAGGGATATATGATGCGATG        2220

      N   K   G   R   M   L   A   K   G   D   L   I   H   Y   L   N   A   G   D   S
     AATAAGGGCCGAATGTTGGCCAAAGGCGACTTAATACATTATTTAAACGCCGGCGATAGC        2280

      V   I   G   D   I   Y   K   N   I   K   E   P   C   L   I   K   V   G   L   F
     GTAATTGGAGATATATATAAAAATATCAAAGAGCCATGTTTGATTAAAGTTGGCCTTTTC        2340

      E   N   D   K   L   L   G   F   S   S   I   T   H   S   N   T   G   Y   C   H
     GAAAATGATAAACTTCTGGGATTTTCTTCTATAACCCATTCAAATACAGGGTATTGTCAT        2400
```

Figure 8/2

```
     Q  G  V  I  F  P  K  N  H  S  E  Y  D  L  R  Y  K  I  C  A
     CAAGGGGTGATTTTCCCAAAGAATCATTCAGAATATGATCTAAGGTATAAAATATGTGCT      2460

     D  Y  K  L  I  Q  E  V  F  P  E  G  L  R  S  L  S  L  I  T
     GATTATAAGCTTATTCAAGAGGTGTTTCCTGAAGGGTTAAGATCTCTATCTTTGATTACT      2520

     S  G  Y  V  K  Y  D  M  G  G  V  S  S  K  K  R  I  L  R  D
     TCGGGTTATGTAAAATATGATATGGGGGGAGTATCTTCAAAAAAAAGAATTTTAAGAGAT      2580

     K  E  L  A  K  I  M  F  E  K  N  K  K  N  L  I  K  F  I  P
     AAAGAGCTTGCCAAAATTATGTTTGAAAAAAATAAAAAAAAACCTTATTAAGTTTATTCCA      2640

     I  S  I  I  K  I  L  F  P  E  R  L  R  R  V  L  R  K  M  Q
     ATTTCAATAATCAAAATTTTATTCCCTGAACGTTTAAGAAGAGTATTGCGGAAAATGCAA      2700


                                       Start of orf4    End of orf3
     Y  I  C  L  T  L  F  F  M  K  N  S  S  P  Y  D  N  E  *
                                            M  I  M  N  K  I
     TATATTTGTCTAACTTTATTCTTCATGAAGAATAGTTCACCATATGATAATGAATAAAAT      2760

     K  K  I  L  K  F  C  T  L  K  K  Y  D  T  S  S  A  L  G  R
     CAAAAAAATACTTAAATTTTGCACTTTAAAAAAAATATGATACATCAAGTGCTTTAGGTAG      2820

     E  Q  E  R  Y  R  I  I  S  L  S  V  I  S  S  L  I  S  K  I
     AGAACAGGAAAGGTACAGGATTATATCCTTGTCTGTTATTTCAAGTTTGATTAGTAAAAT      2880

     L  S  L  L  S  L  I  L  T  V  S  L  T  L  P  Y  L  G  Q  E
     ACTCTCACTACTTTCTCTTATATTAACTGTAAGTTTAACTTTACCTTATTTAGGACAAGA      2940

     R  F  G  V  W  M  T  I  T  S  L  G  A  A  L  T  F  L  D  L
     GAGATTTGGTGTATGGATGACTATTACCAGTCTTGGTGCTGCTCTGACATTTTTGGACTT      3000

     G  I  G  N  A  L  T  N  R  I  A  H  S  F  A  C  G  K  N  L
     AGGTATAGGAAATGCATTAACAAACAGGATCGCACATTCATTTGCGTGTGGCAAAAATTT      3060

     K  M  S  R  Q  I  S  G  G  L  T  L  L  A  G  L  S  F  V  I
     AAAGATGAGTCGGCAAATTAGTGGTGGGCTCACTTTGCTGGCTGGATTATCGTTTGTCAT      3120

     T  A  I  C  Y  I  T  S  G  M  I  D  W  Q  L  V  I  K  G  I
     AACTGCAATATGCTATATTACTTCTGGCATGATTGATTGGCAACTAGTAATAAAAGGTAT      3180

     N  E  N  V  Y  A  E  L  Q  H  S  I  K  V  F  V  I  I  F  G
     AAACGAGAATGTGTATGCAGAGTTACAACACTCAATTAAAGTCTTTGTAATCATATTTGG      3240

     L  G  I  Y  S  N  G  V  Q  K  V  Y  M  G  I  Q  K  A  Y  I
     ACTTGGAATTTATTCAAATGGTGTGCAAAAAGTTTATATGGGAATACAAAAAGCCTATAT      3300

     S  N  I  V  N  A  I  F  I  L  L  S  I  I  T  L  V  I  S  S
     AAGTAATATTGTTAATGCCATATTTATATTGTTATCTATTATTACTCTAGTAATATCGTC      3360

     K  L  H  A  G  L  P  V  L  I  V  S  T  L  G  I  Q  Y  I  S
     GAAACTACATGCGGGACTACCAGTTTTAATTGTCAGCACTCTTGGTATTCAATACATATC      3420

     G  I  Y  L  T  I  N  L  I  I  K  R  L  I  K  F  T  K  V  N
     GGGAATCTATTTAACAATTAATCTTATTATAAAGCGATTAATAAAGTTTACAAAGTTAA      3480

     I  H  A  K  R  E  A  P  Y  L  I  L  N  G  F  F  F  F  I  L
     CATACATGCTAAAAGAGAAGCTCCATATTTGATATTAAACGGTTTTTTCTTTTTTATTTT      3540

     Q  L  G  T  L  A  T  W  S  G  D  N  F  I  I  S  I  T  L  G
     ACAGTTAGGCACTCTGGCAACATGGAGTGGTGATAACTTTATAATATCTATAACATTGGG      3600
```

Figure 8/3

```
      V  T  Y  V  A  V  F  S  I  T  Q  R  L  F  Q  I  S  T  V  P
    TGTTACTTATGTTGCTGTTTTTAGCATTACACAGAGATTATTTCAAATATCTACGGTCCC      3660

         L  T  I  Y  N  I  P  L  W  A  A  Y  A  D  A  H  A  R  N  D
    TCTTACGATTTATAACATCCCGTTATGGGCTGCTTATGCAGATGCTCATGCACGCAATGA      3720

         T  Q  F  I  K  K  T  L  R  T  S  L  K  I  V  G  I  S  S  F
    TACTCAATTTATAAAAAAGACGCTCAGAACATCATTGAAAATAGTGGGTATTTCATCATT      3780

         L  L  A  F  I  L  V  V  F  G  S  E  V  V  N  I  W  T  E  G
    CTTATTGGCCTTCATATTAGTAGTGTTCGGTAGTGAAGTCGTTAATATTTGGACAGAAGG      3840

         K  I  Q  V  P  R  T  F  I  I  A  Y  A  L  W  S  V  I  D  A
    AAAGATTCAGGTACCTCGAACATTCATAATAGCTTATGCTTTATGGTCTGTTATTGATGC      3900

         F  S  N  T  F  A  S  F  L  N  G  L  N  I  V  K  Q  Q  M  L
    TTTTTCGAATACATTTGCAAGCTTTTTAAATGGTTTGAACATAGTTAAACAACAAATGCT      3960

         A  V  V  T  L  I  L  I  A  I  P  A  K  Y  I  I  V  S  H  F
    TGCTGTTGTAACATTGATATTGATCGCAATTCCAGCAAAATACATCATAGTTAGCCATTT      4020

         G  L  T  V  M  L  Y  C  F  I  F  I  Y  I  V  N  Y  F  I  W
    TGGGTTAACTGTTATGTTGTACTGCTTCATTTTTATATATATTGTAAATTACTTTATATG      4080
```

Start of orf5, End of orf4

```
                                                   M  K  M
         Y  K  C  S  F  K  K  H  I  D  R  Q  L  N  I  R  G  *
    GTATAAATGTAGTTTTAAAAAACATATCGATAGACAGTTAAATATAAGAGG ATG AAAATG      4140

         K  Y  I  P  V  Y  Q  P  S  L  T  G  K  E  K  E  Y  V  N  E
    AAATATATACCAGTTTACCAACCGTCATTGACAGGAAAAGAAAAAGAATATGTAAATGAA      4200

         C  L  D  S  T  W  I  S  S  K  G  N  Y  I  Q  K  F  E  N  K
    TGTCTGGACTCAACGTGGATTTCATCAAAAGGAAACTATATTCAGAAGTTTGAAAATAAA      4260

         F  A  E  Q  N  H  V  Q  Y  A  T  T  V  S  N  G  T  V  A  L
    TTTGCGGAACAAAACCATGTGCAATATGCAACTACTGTAAGTAATGGAACGGTTGCTCTT      4320

         H  L  A  L  L  A  L  G  I  S  E  G  D  E  V  I  V  P  T  L
    CATTTAGCTTTGTTAGCGTTAGGTATATCGGAAGGAGATGAAGTTATTGTTCCAACACTG      4380

         T  Y  I  A  S  V  N  A  I  K  Y  T  G  A  T  P  I  F  V  D
    ACATATATAGCATCAGTTAATGCTATAAAATACACAGGAGCCACCCCCATTTTCGTTGAT      4440

         S  D  N  E  T  W  Q  M  S  V  S  D  I  E  Q  K  I  T  N  K
    TCAGATAATGAAACTTGGCAAATGTCTGTTAGTGACATAGAACAAAAAATCACTAATAAA      4500

         T  K  A  I  M  C  V  H  L  Y  G  H  P  C  D  M  E  Q  I  V
    ACTAAAGCTATTATGTGTGTGTCCATTTATACGGACATCCATGTGATATGGAACAAATTGTA      4560

         E  L  A  K  S  R  N  L  F  V  I  E  D  C  A  E  A  F  G  S
    GAACTGGCCAAAAGTAGAAATTTGTTTGTAATTGAAGATTGCGCTGAAGCCTTTGGTTCT      4620

         K  Y  K  G  K  Y  V  G  T  F  G  D  I  S  T  F  S  F  F  G
    AAATATAAAGGTAAATATGTGGGAACATTTGGAGATATTTCTACTTTTAGCTTTTTTGGA      4680

         N  K  T  I  T  T  G  E  G  G  M  V  V  T  N  D  K  T  L  Y
    AATAAAACTATTACTACAGGTGAAGGTGGAATGGTTGTCACGAATGACAAAACACTTTAT      4740

         D  R  C  L  H  F  K  G  Q  G  L  A  V  H  R  Q  Y  W  H  D
    GACCGTTGTTTACATTTTAAAGGCCAAGGATTAGCTGTACATAGGCAATATTGGCATGAC      4800

         V  I  G  Y  N  Y  R  M  T  N  I  C  A  A  I  G  L  A  Q  L
    GTTATAGGCTACAATTATAGGATGACAAATATCTGCGCTGCTATAGGATTAGCCCAGTTA      4860
```

Figure 8/4

09/423093

```
    E   Q   A   D   D   F   I   S   R   K   R   E   I   A   D   I   Y   K   K   N
    GAACAAGCTGATGATTTTATATCACGAAAACGTGAAATTGCTGATATTTATAAAAAAAT         4920

    I   N   S   L   V   Q   V   H   K   E   S   K   D   V   F   H   T   Y   W   M
    ATCAACAGTCTTGTACAAGTCCACAAGGAAAGTAAAGATGTTTTTCACACTTATTGGATG         4980

    V   S   I   L   T   R   T   A   E   E   R   E   E   L   R   N   H   L   A   D
    GTCTCAATTCTAACTAGGACCGCAGAGGAAAGAGAGGAATTAAGGAATCACCTTGCAGAT         5040

    K   L   I   E   T   R   P   V   F   Y   P   V   H   T   M   P   M   Y   S   E
    AAACTCATCGAAACAAGGCCAGTTTTTTACCCTGTCCACACGATGCCAATGTACTCGGAA         5100

    K   Y   Q   K   H   P   I   A   E   D   L   G   W   R   G   I   N   L   P   S
    AAATATCAAAAGCACCCTATAGCTGAGGATCTTGGTTGGCGTGGAATTAATTTACCTAGT         5160

    F   P   S   L   S   N   E   Q   V   I   Y   I   C   E   S   I   N   E   F   Y
    TTCCCCAGCCTATCGAATGAGCAAGTTATTTATATTTGTGAATCTATTAACGAATTTTAT         5220


                End of orf5                        Start of orf6
    S   D   K   *                                  M   K   I   A   L   N   S   D
    AGTGATAAATAGCCTAAAATATTGTAAAGGTCATTCATGAAAATTGCGTTGAATTCAGAT         5280

    G   F   Y   E   W   G   G   G   I   D   F   I   K   Y   I   L   S   I   L   E
    GGATTTTACGAGTGGGGCGGTGGAATTGATTTTATTAAATATATTCTGTCAATATTAGAA         5340

    T   K   P   E   I   C   I   D   I   L   L   P   R   N   D   I   H   S   L   I
    ACGAAACCAGAAATATGTATCGATATTCTTTTACCGAGAAATGATATACATTCTCTTATA         5400

    R   E   K   A   F   P   F   K   S   I   L   K   A   I   L   K   R   E   R   P
    AGAGAAAAAGCATTTCCTTTTAAAAGTATATTAAAAGCAATTTTAAAGAGGGAAAGGCCT         5460

    R   W   I   S   L   N   R   F   N   E   Q   Y   Y   R   D   A   F   T   Q   N
    CGATGGATTTCATTAAATAGATTTAATGAGCAATACTATAGAGATGCCTTTACACAAAAT         5520

    N   I   E   T   N   L   T   F   I   K   S   K   S   S   A   F   Y   S   Y   F
    AATATAGAGACGAATCTTACCTTTATTAAAAGTAAGAGCTCTGCCTTTTATTCATATTTT         5580

    D   S   S   D   C   D   V   I   L   P   C   M   R   V   P   S   G   N   L   N
    GATAGTAGCGATTGTGATGTTATTCTTCCTTGCATGCGTGTTCCTTCGGGAAATTTGAAT         5640

    K   K   A   W   I   G   Y   I   Y   D   F   Q   H   C   Y   Y   P   S   F   F
    AAAAAAGCATGGATTGGTTATATTTATGACTTTCAACACTGTTACTATCCTTCATTTTTT         5700

    S   K   R   E   I   D   Q   R   N   V   F   F   K   L   M   L   N   C   A   N
    AGTAAGCGAGAAATAGATCAAAGGAATGTGTTTTTTAAATTGATGCTCAATTGCGCTAAC         5760

    N   I   I   V   N   A   H   S   V   I   T   D   A   N   K   Y   V   G   N   Y
    AATATTATTGTTAATGCACATTCAGTTATTACCGATGCAAATAAATATGTTGGGAATTAT         5820

    S   A   K   L   H   S   L   P   F   S   P   C   P   Q   L   K   W   F   A   D
    TCTGCAAAACTACATTCTCTTCCATTTAGTCCATGCCCTCAATTAAAATGGTTCGCTGAT         5880

    Y   S   G   N   I   A   K   Y   N   I   D   K   D   Y   F   I   I   C   N   Q
    TACTCTGGTAATATTGCCAAATATAATATTGACAAGGATTATTTTATAATTTGCAATCAA         5940

    F   W   K   H   K   D   H   A   T   A   F   R   A   F   K   I   Y   T   E   Y
    TTTTGGAAACATAAAGATCATGCAACTGCTTTTAGGGCATTTAAAATTTATACTGAATAT         6000

    N   P   D   V   Y   L   V   C   T   G   A   T   Q   D   Y   R   F   P   G   Y
    AATCCTGATGTTTATTTAGTATGCACGGGAGCTACTCAAGATTATCGATTCCCTGGATAT         6060

    F   N   E   L   M   V   L   A   K   K   L   G   I   E   S   K   I   K   I   L
    TTTAATGAATTGATGGTTTTTGGCAAAAAAGCTCGGAATTGAATCGAAAATTAAGATATTA         6120
```

Figure 8/5

```
      G  H  I  P  K  L  E  Q  I  E  L  I  K  N  C  I  A  V  I  Q
      GGGCATATACCTAAACTTGAACAAATTGAATTAATCAAAAATTGCATTGCTGTAATACAA       6180

      P  T  L  F  E  G  G  P  G  G  G  V  T  F  D  A  I  A  L  G
      CCAACCTTATTTGAAGGCGGGCCTGGAGGGGGGGTAACATTTGACGCTATTGCATTAGGG       6240

      K  K  V  I  L  S  D  I  D  V  N  K  E  V  N  C  G  D  V  Y
      AAAAAAGTTATACTATCTGACATAGATGTCAATAAAGAAGTTAATTGCGGTGATGTATAT       6300

      F  F  Q  A  K  N  H  Y  S  L  N  D  A  M  V  K  A  D  E  S
      TTCTTTCAGGCAAAAAACCATTATTCATTAAATGACGCGATGGTAAAAGCTGATGAATCT       6360

      K  I  F  Y  E  P  T  T  L  I  E  L  G  L  K  R  R  N  A  C
      AAAATTTTTTATGAACCTACAACTCTGATAGAATTGGGTCTCAAAAGACGCAATGCGTGT       6420
```

```
                                                         End of orf6
      A  D  F  L  L  D  V  V  K  Q  E  I  E  S  R  S       *
      GCAGATTTTCTTTTAGATGTTGTGAAACAAGAAATTGAATCCCGATCT TAATATATTCAA       6480
```

```
          Start of orf7
          M  T  K  V  A  L  I  T  G  V  T  G  Q  D  G  S  Y
      GAGGTATATAATGACTAAAGTCGCTCTTATTACAGGTGTAACTGGACAAGATGGATCTTA       6540

      L  A  E  F  L  L  D  K  G  Y  E  V  H  G  I  K  R  R  A  S
      TCTAGCTGAGTTTTTGCTTGATAAAGGGTATGAAGTTCATGGTATCAAACGCCGAGCCTC       6600

      S  F  N  T  E  R  I  D  H  I  Y  Q  D  P  H  G  S  N  P  N
      ATCTTTTAATACAGAACGCATAGACCATATTTATCAAGATCCACATGGTTCTAACCCAAA       6660

      F  H  L  H  Y  G  D  L  T  D  S  S  N  L  T  R  I  L  K  E
      TTTTCACTTGCACTATGGAGATCTGACTGATTCATCTAACCTCACTAGAATTCTAAAGGA       6720

      V  Q  P  D  E  V  Y  N  L  A  A  M  S  H  V  A  V  S  F  E
      GGTACAGCCAGATGAAGTATATAATTTAGCTGCTATGAGTCACGTAGCAGTTTCTTTTGA       6780

      S  P  E  Y  T  A  D  V  D  A  I  G  T  L  R  L  L  E  A  I
      GTCTCCAGAATATACAGCCGATGTCGATGCAATTGGTACATTACGTTTACTGGAAGCAAT       6840

      R  F  L  G  L  E  N  K  T  R  F  Y  Q  A  S  T  S  E  L  Y
      TCGCTTTTTAGGATTGGAAAACAAAACGCGTTTCTATCAAGCTTCAACCTCAGAATTATA       6900

      G  L  V  Q  E  I  P  Q  K  E  S  T  P  F  Y  P  R  S  P  Y
      TGGACTTGTTCAGGAAATCCCTCAAAAAGAATCCACCCCTTTTTATCCTCGTTCCCCTTA       6960

      A  V  A  K  L  Y  A  Y  W  I  T  V  N  Y  R  E  S  Y  G  I
      TGCAGTTGCAAAACTTTACGCATATTGGATCACGGTAAATTATCGAGAGTCATATGGTAT       7020

      Y  A  C  N  G  I  L  F  N  H  E  S  P  R  R  G  E  T  F  V
      TTATGCATGTAATGGTATATTGTTCAATCATGAATCTCCACGCCGTGGAGAAACGTTTGT       7080

      T  R  K  I  T  R  G  L  A  N  I  A  Q  G  L  E  S  C  L  Y
      AACAAGGAAAATTACTCGAGGACTTGCAAATATTGCACAAGGCTTGGAATCATGTTTGTA       7140

      L  G  N  M  D  S  L  R  D  W  G  H  A  K  D  Y  V  R  M  Q
      TTTAGGGAATATGGATTCGTTACGAGATTGGGGACATGCAAAAGATTATGTTAGAATGCA       7200

      W  L  M  L  Q  Q  E  Q  P  E  D  F  V  I  A  T  G  V  Q  Y
      ATGGTTGATGTTACAACAGGAGCAACCCGAAGATTTTGTGATTGCAACAGGAGTCCAATA       7260

      S  V  R  Q  F  V  E  M  A  A  A  Q  L  G  I  K  M  S  F  V
      CTCAGTCCGTCAGTTTGTCGAAATGGCAGCAGCACAACTTGGTATTAAGATGAGCTTTGT       7320
```

Figure 8/6

09/423093

```
         G   K   G   I   E   E   K   G   I   V   D   S   V   E   G   Q   D   A   P   G
        TGGTAAAGGAATCGAAGAAAAAGGCATTGTAGATTCGGTTGAAGGACAGGATGCTCCAGG        7380

           V   K   P   G   D   V   I   V   A   V   D   P   R   Y   F   R   P   A   E   V
        TGTGAAACCAGGTGATGTCATTGTTGCTGTTGATCCTCGTTATTTCCGACCAGCTGAAGT        7440

           D   T   L   L   G   D   P   S   K   A   N   L   K   L   G   W   R   P   E   I
        TGATACTTTGCTTGGAGATCCGAGCAAAGCTAATCTCAAACTTGGTTGGAGACCAGAAAT        7500

             T   L   A   E   M   I   S   E   M   V   A   K   D   L   E   A   A   K   K   H
        TACTCTTGCTGAAATGATTTCTGAAATGGTTGCCAAAGATCTTGAAGCCGCTAAAAAACA        7560
```

                                                     Start of orf8, End of orf7
                                                        M   M   M   N   K
```
         S   L   L   K   S   H   G   F   S   V   S   L   A   L   E   *
        TTCTCTTTTAAAATCGCATGGTTTTTCTGTAAGCTTAGCTCTGGAATGATGATGAATAAG        7620

           Q   R   I   F   I   A   G   H   Q   G   M   V   G   S   A   I   T   R   R   L
        CAACGTATTTTTATTGCTGGTCACCAAGGAATGGTTGGATCAGCTATTACCCGACGCCTC        7680

         K   Q   R   D   D   V   E   L   V   L   R   T   R   D   E   L   N   L   L   D
        AAACAACGTGATGATGTTGAGTTGGTTTTACGTACTCGGGATGAATTGAACTTGTTGGAT        7740

         S   S   A   V   L   D   F   F   S   S   Q   K   I   D   Q   V   Y   L   A   A
        AGTAGCGCTGTTTTGGATTTTTTTTTCTTCACAGAAAATCGACCAGGTTTATTTGGCAGCA        7800

         A   K   V   G   G   I   L   A   N   S   S   Y   P   A   D   F   I   Y   E   N
        GCAAAAGTCGGAGGTATTTTAGCTAACAGTTCTTATCCTGCCGATTTTATATATGAGAAT        7860

         I   M   I   E   A   N   V   I   H   A   A   H   K   N   N   V   N   K   L   L
        ATAATGATAGAGGCGAATGTCATTCATGCTGCCCACAAAAATAATGTAAATAAACTGCTT        7920

         F   L   G   S   S   C   I   Y   P   K   L   A   H   Q   P   I   M   E   D   E
        TTCCTCGGTTCGTCGTGTATTTATCCTAAGTTAGCACACCAACCGATTATGGAAGACGAA        7980

         L   L   Q   G   K   L   E   P   T   N   E   P   Y   A   I   A   K   I   A   G
        TTATTACAAGGGAAACTTGAGCCAACAAATGAACCTTATGCTATCGCAAAAATTGCAGGT        8040

         I   K   L   C   E   S   Y   N   R   Q   F   G   R   D   Y   R   S   V   M   P
        ATTAAATTATGTGAATCTTATAACCGTCAGTTTGGGCGTGATTACCGTTCAGTAATGCCA        8100

         T   N   L   Y   G   P   N   D   N   F   H   P   S   N   S   H   V   I   P   A
        ACCAATCTTTATGGTCCAAATGACAATTTTCATCCAAGTAATTCTCATGTGATTCCGGCG        8160

         L   L   R   R   F   H   D   A   V   E   N   N   S   P   N   V   V   V   W   G
        CTTTTGCGCCGCTTTCATGATGCTGTGGAAAACAATTCTCCGAATGTTGTTGTTTGGGGA        8220

         S   G   T   P   K   R   E   F   L   H   V   D   D   M   A   S   A   S   I   Y
        AGTGGTACTCCAAAGCGTGAATTCTTACATGTAGATGATATGGCTTCTGCAAGCATTTAT        8280

         V   M   E   M   P   Y   D   I   W   Q   K   N   T   K   V   M   L   S   H   I
        GTCATGGAGATGCCATACGATATATGGCAAAAAAATACTAAAGTAATGTTGTCTCATATC        8340

         N   I   G   T   G   I   D   C   T   I   C   E   L   A   E   T   I   A   K   V
        AATATTGGAACAGGTATTGACTGCACGATTTGTGAGCTTGCGGAAACAATAGCAAAAGTT        8400

         V   G   Y   K   G   H   I   T   F   D   T   T   K   P   D   G   A   P   R   K
        GTAGGTTATAAAGGGCATATTACGTTCGATACAACAAAGCCCGATGGAGCCCCTCGAAAA        8460

         L   L   D   V   T   L   L   H   Q   L   G   W   N   H   K   I   T   L   H   K
        CTACTTGATGTAACGCTTCTTCATCAACTAGGTTGGAATCATAAAATTACCCTTCACAAG        8520
```

Figure 8/7

```
                                                            End of orf8
        G   L   E   N   T   Y   N   W   F   L   E   N   Q   L   Q   Y   R   G   *
        GGTCTTGAAAATACATACAACTGGTTTCTTGAAAACCAACTTCAATATCGGGGG TAATAA        8580

Start of orf9
        M   F   L   H   S   Q   D   F   A   T   I   V   R   S   T   P   L   I   S   I
        TGTTTTTACATTCCCAAGACTTTGCCACAATTGTAAGGTCTACTCCTCTTATTTCTATAG        8640

        D   L   I   V   E   N   E   F   G   E   I   L   L   G   K   R   I   N   R   P
        ATTTGATTGTGGAAAACGAGTTTGGCGAAATTTTGCTAGGAAAACGAATCAACCGCCCGG        8700

        A   Q   G   Y   W   F   V   P   G   G   R   V   L   K   D   E   K   L   Q   T
        CACAGGGCTATTGGTTCGTTCCTGGTGGTAGGGTGTTGAAAGATGAAAAATTGCAGACAG        8760

        A   F   E   R   L   T   E   I   E   L   G   I   R   L   P   L   S   V   G   K
        CCTTTGAACGATTGACAGAAATTGAACTAGGAATTCGTTTGCCTCTCTCTGTGGGTAAGT        8820

        F   Y   G   I   W   Q   H   F   Y   E   D   N   S   M   G   G   D   F   S   T
        TTTATGGTATCTGGCAGCACTTCTACGAAGACAATAGTATGGGGGGAGACTTTTCAACGC        8880

        H   Y   I   V   I   A   F   L   L   K   L   Q   P   N   I   L   K   L   P   K
        ATTATATAGTTATAGCATTCCTTCTTAAATTACAACCAAACATTTTGAAATTACCGAAGT        8940

        S   Q   H   N   A   Y   C   W   L   S   R   A   K   L   I   N   D   D   D   V
        CACAACATAATGCTTATTGCTGGCTATCGCGAGCAAAGCTGATAAATGATGACGATGTGC        9000

        H   Y   N   C   R   A   Y   F   N   N   K   T   N   D   A   I   G   L   D   N
        ATTATAATTGTCGCGCATATTTTAACAATAAAACAAATGATGCGATTGGCTTAGATAATA        9060


                Start of orf10      End of orf9
                    M   S   D   A   P   I   I   A   V   V   M   A   G   G   T   G   S
        K   D   I   I   C   L   M   R   Q   *
        AGGATATAATATGTCTGATGCGCCAA TAATTGCTGTAGTTATGGCCGGTGGTACAGGCAG        9120

            R   L   W   P   L   S   R   E   L   Y   P   K   Q   F   L   Q   L   S   G   D
        TCGTCTTTGGCCACTTTCTCGTGAACTATATCCAAAGCAGTTTTTACAACTCTCTGGTGA        9180

            N   T   L   L   Q   T   T   L   L   R   L   S   G   L   S   C   Q   K   P   L
        TAACACCTTGTTACAAACGACTTTGCTACGACTTTCAGGCCTATCATGTCAAAAACCATT        9240

            V   I   T   N   E   Q   H   R   F   V   V   A   E   Q   L   R   E   I   N   K
        AGTGATAACAAATGAACAGCATCGCTTTGTTGTGGCTGAACAGTTAAGGGAAATAAATAA        9300

            L   N   G   N   I   I   L   E   P   C   G   R   N   T   A   P   A   I   A   I
        ATTAAATGGTAATATTATTCTAGAACCATGCGGGCGAAATACTGCACCAGCAATAGCGAT        9360

            S   A   F   H   A   L   K   R   N   P   Q   E   D   P   L   L   L   V   L   A
        ATCTGCGTTTCATGCGTTAAAACGTAATCCTCAGGAAGATCCATTGCTTCTAGTTCTTGC        9420

            A   D   H   V   I   A   K   E   S   V   F   C   D   A   I   K   N   A   T   P
        GGCAGACCACGTTATAGCTAAAGAAAGTGTTTTCTGTGATGCTATTAAAAATGCAACTCC        9480

            I   A   N   Q   G   K   I   V   T   F   G   I   I   P   E   Y   A   E   T   G
        CATCGCTAATCAAGGTAAAATTGTAACGTTTGGAATTATACCAGAATATGCTGAAACTGG        9540

            Y   G   Y   I   E   R   G   E   L   S   V   P   L   Q   G   H   E   N   T   G
        TTATGGGTATATTGAGAGAGGTGAACTATCTGTACCGCTTCAAGGGCATGAAAATACTGG        9600

            F   Y   Y   V   N   K   F   V   E   K   P   N   R   E   T   A   E   L   Y   M
        TTTTTATTATGTAAATAAGTTTGTCGAAAAGCCTAATCGTGAAACCGCAGAATTGTATAT        9660

            T   S   G   N   H   Y   W   N   S   G   I   F   M   F   K   A   S   V   Y   L
        GACTTCTGGTAATCACTATTGGAATAGTGGAATATTCATGTTTAAGGCATCTGTTTATCT        9720
```

Figure 8/8

```
         E    E    L    R    K    F    R    P    D    I    Y    N    V    C    E    Q    V    A    S    S
        TGAGGAATTGAGAAAATTTAGACCTGACATTTACAATGTTTGTGAACAGGTTGCCTCATC                9780

         S    Y    I    D    L    D    F    I    R    L    S    K    E    Q    F    Q    D    C    P    A
        CTCATACATTGATCTAGATTTTATTCGATTATCAAAAGAACAATTTCAAGATTGTCCTGC                9840

         E    S    I    D    F    A    V    M    E    K    T    E    K    C    V    V    C    P    V    D
        TGAATCTATTGATTTTGCTGTAATGGAAAAAACAGAAAAATGTGTTGTATGCCCTGTTGA                9900

         I    G    W    S    D    V    G    S    W    Q    S    L    W    D    I    S    L    K    S    K
        TATTGGTTGGAGTGACGTTGGATCTTGGCAATCGTTATGGGACATTAGTCTAAAATCGAA                9960

         T    G    D    V    C    K    G    D    I    L    T    Y    D    T    K    N    N    Y    I    Y
        AACAGGAGATGTATGTAAAGGTGATATATTAACCTATGATACTAAGAATAATTATATCTA                10020

         S    E    S    A    L    V    A    A    I    G    I    E    D    M    V    I    V    Q    T    K
        CTCTGAGTCAGCGTTGGTAGCCGCCATTGGAATTGAAGATATGGTTATCGTGCAAACTAA                10080

         D    A    V    L    V    S    K    K    S    D    V    Q    H    V    K    K    I    V    E    M
        AGATGCCGTTCTTGTGTCTAAAAAGAGTGATGTACAGCATGTAAAAAAAATAGTCGAAAT                10140

         L    K    L    Q    Q    R    T    E    Y    I    S    H    R    E    V    F    R    P    W    G
        GCTTAAATTGCAGCAACGTACAGAGTATATTAGTCATCGTGAAGTTTTCCGACCATGGGG                10200

         K    F    D    S    I    D    Q    G    E    R    Y    K    V    K    K    I    I    V    K    P
        AAAATTTGATTCGATTGACCAAGGTGAGCGATACAAAGTCAAGAAAATTATTGTGAAACC                10260

         G    E    G    L    S    L    R    M    H    H    H    R    S    E    H    W    I    V    L    S
        TGGTGAGGGGCTTTCTTTAAGGATGCATCACCATCGTTCTGAACATTGGATCGTGCTTTC                10320

         G    T    A    K    V    T    L    G    D    K    T    K    L    V    T    A    N    E    S    I
        TGGTACAGCAAAAGTAACCCTTGGCGATAAAACTAAACTAGTCACCGCAAATGAATCGAT                10380

         Y    I    P    L    G    A    A    Y    S    L    E    N    P    G    I    I    P    L    N    L
        ATACATTCCCCTTGGCGCAGCGTATAGTCTTGAGAATCCGGGCATAATCCCTCTTAATCT                10440

         I    E    V    S    S    G    D    Y    L    G    E    D    D    I    I    R    Q    K    E    R
        TATTGAAGTCAGTTCAGGGGATTATTTGGGAGAGGATGATATTATAAGACAGAAAGAACG                10500

                  End of orf10    Start of orf11
         Y    K    H    E    D    *         M    K    S    L    T    C    F    K    A    Y    D    I    R
        TTACAAACATGAAGAT TAACATATGAAATCTTTAACCTGCTTTAAAGCCTATGATATTCG                10560

         G    K    L    G    E    E    L    N    E    D    I    A    W    R    I    G    R    A    Y    G
        CGGGAAATTAGGCGAAGAACTGAATGAAGATATTGCCTGGCGCATTGGGCGTGCCTATGG                10620

         E    F    L    K    P    K    T    I    V    L    G    G    D    V    R    L    T    S    E    A
        CGAATTTCTCAAACCGAAAACCATTGTTTTAGGCGGTGATGTCCGCCTCACCAGCGAAGC                10680

         L    K    L    A    L    A    K    G    L    Q    D    A    G    V    D    V    L    D    I    G
        GTTAAAACTGGCGCTTGCGAAAGGTTTACAGGATGCGGGCGTCGATGTGCTGGATATCGG                10740

         M    S    G    T    E    E    I    Y    F    A    T    F    H    L    G    V    D    G    G    I
        TATGTCCGGCACCGAAGAGATCTATTTCGCCACGTTCCATCTCGGAGTGGATGGCGGCAT                10800

         E    V    T    A    S    H    N    P    M    D    Y    N    G    M    K    L    V    R    E    G
        CGAAGTTACCGCCAGCCATAACCCGATGGATTACAACGGCATGAAGCTGGTGCGCGAAGG                10860

         A    R    P    I    S    G    D    T    G    L    R    D    V    Q    R    L    A    E    A    N
        GGCTCGCCCGATCAGCGGTGATACCGGACTGCGCGATGTCCAGCGTCTGGCAGAAGCCAA                10920

         D    F    P    P    V    D    E    T    K    R    G    R    Y    Q    Q    I    N    L    R    D
        TGACTTCCCTCCTGTCGATGAAACCAAACGTGGTCGCTATCAGCAAATCAATCTGCGTGA                10980
```

Figure 8/9

```
      A   Y   V   D   H   L   F   G   Y   I   N   V   K   N   L   T   P   L   K   L
    CGCTTACGTTGATCACCTGTTCGGTTATATCAACGTCAAAAACCTCACGCCGCTCAAGCT          11040

      V   I   N   S   G   N   G   A   A   G   P   V   V   D   A   I   E   A   R   F
    GGTGATCAACTCCGGGAACGGCGCAGCGGGTCCGGTGGTGGACGCCATTGAAGCCCGATT          11100

      K   A   L   G   A   P   V   E   L   I   K   V   H   N   T   P   D   G   N   F
    TAAAGCCCTCGGCGCACCGGTGGAATTAATCAAAGTACACAACACGCCGGACGGCAATTT          11160

      P   N   G   I   P   N   P   L   L   P   E   C   R   D   D   T   R   N   A   V
    CCCCAACGGTATTCCTAACCCGCTGCTGCCGGAATGCCGCGACGACACCCGTAATGCGGT          11220

      I   K   H   G   A   D   M   G   I   A   F   D   G   D   F   D   R   C   F   L
    CATCAAACACGGCGCGGATATGGGCATTGCCTTTGATGGCGATTTTGACCGCTGTTTCCT          11280

      F   D   E   K   G   Q   F   I   E   G   Y   Y   I   V   G   L   L   A   E   A
    GTTTGACGAAAAAGGGCAGTTTATCGAGGGCTACTACATTGTCGGCCTGCTGGCAGAAGC          11340

      F   L   E   K   N   P   G   A   K   I   I   H   D   P   R   L   S   W   N   T
    GTTCCTCGAAAAAAATCCCGGCGCGAAGATCATCCACGATCCACGTCTCTCCTGGAACAC          11400

      V   D   V   V   T   A   A   G   G   T   P   V   M   S   K   T   G   H   A   F
    CGTTGATGTGGTGACTGCCGCAGGCGGCACCCCGGTAATGTCGAAAACCGGACACGCCTT          11460

      I   K   E   R   M   R   K   E   D   A   I   Y   G   G   E   M   S   A   H   H
    TATTAAAGAACGTATGCGCAAGGAAGACGCCATCTACGGTGGCGAAATGAGCGCTCACCA          11520

      Y   F   R   D   F   A   Y   C   D   S   G   M   I   P   W   L   L   V   A   E
    TTACTTCCGTGATTTCGCTTACTGCGACAGCGGCATGATCCCGTGGCTGCTGGTCGCCGA          11580

      L   V   C   L   K   G   K   T   L   G   E   M   V   R   D   R   M   A   A   F
    ACTGGTGTGCCTGAAAGGAAAAACGCTGGGCGAAATGGTGCGCGACCGGATGGCGGCGTT          11640

      P   A   S   G   E   I   N   S   K   L   A   Q   P   V   E   A   I   N   R   V
    TCCGGCAAGCGGTGAGATCAACAGCAAACTGGCGCAACCCGTTGAGGCAATTAATCGCGT          11700

      E   Q   H   F   S   R   E   A   L   A   V   D   R   T   D   G   I   S   M   T
    GGAACAGCATTTTAGCCGCGAGGCGCTGGCGGTGGATCGCACCGATGGCATCAGCATGAC          11760

      F   A   D   W   R   F   N   L   R   S   S   N   T   E   P   V   V   R   L   N
    CTTTGCCGACTGGCGCTTTAACCTGCGCTCCTCCAACACCGAACCGGTGGTGCGGTTGAA          11820

      V   E   S   R   G   D   V   K   L   M   E   K   K   T   K   A   L   L   K   L
    TGTGGAATCACGCGGTGATGTAAAGCTAATGGAAAAGAAAACTAAAGCTCTTCTTAAATT          11880

              End of orf11
      L   S   E   *
    GCTAAGTGAGTGATTATTTACATTAATCATTAAGCGTATTTAAGATTATATTAAAGTAAT          11940

    GTTATTGCGGTATATGATGAATATGTGGGCTTTTTTATGTATAACGACTATACCGCAACT          12000


              Start of H-repeat
    TTATCTAGGAAAAGATTAATAGAAATAAAGTTTTGTACTGACCAATTTGCATTTCACGTC          12060

    ACGATTGAGACGTTCCTTTGCTTAAGACATTTTTTCATCGCTTATGTAATAACAAATGTG          12120

    CCTTATATAAAAAGGAGAACAAAATGGAACTTAAAATAATTGAGACAATAGATTTTTATT          12180

    ATCCCTGTTTACGATATTATAGCCAAAGTTGTATCCTGCATCAGTCCTGCAATATTTCAC          12240

    GAGTGCTTTGTTAACTGAATACATGTCTGCCATTTTCCAGATGATAACGACGTCATCGCA          12300

    ATTGATGGTAAAACACTTCGGCACACTTATGACAAGAGTCGTCGCAGAGGAGTGGTTCAT          12360
```

Figure 8/10

```
GTCATTAGTGCGTTTCAGCAATGCACAGTCTGGTCCTCGGATAGATCAAGACGGATGAGA        12420

AACCTAATGCGTTCACAGTTATTCATGAACTTTCTAAAATGATGGGTATTAAAGGAAAAA        12480

TAATCATAACTGATGCGATGGCTTGCCAGAAAGATATTGCAGAGAAGATATAAAAACAGA        12540

GATGTGATTATTTATTCGCTGTAAAAGGAAATAAGAGTCGGCTTAATAGAGTCTTTGAGG        12600

AGATATTTACGCTGAAAGAATTAAATAATCCAAAACATGACAGTTACGCAATTAGTGAAA        12660

AGAGGCACGGCAGAGACGATGTCCGTCTTCATATTGTTTGAGATGCTCCTGATGAGCTTA        12720

TTGATTTCACGTTTGAATGGAAAGGGCTGCAGAATTTATGAATGGCAGTCCACTTTCTCT        12780

CAATAATAGCAGAGCAAAAGAAAGAATCCGAAATGACGATCAAATATTATATTAGATCTG        12840

CTGCTTTAACCGCAGAGAAGTTCGCCACAGTAAATCGAAATCACTGGCGCATGGAGAATA        12900

AGTTGCACAGTAGCCTGATGTGGTAATGAATGAAATCGACTATAATATAAGAAGGCGAGT        12960

TGCATTCGAATGATTTTCTAGAATGCGGCACATCGCTATTAATATCTGACAATGATAATG        13020

TATTCAAGGCAGGATTATCATGTAAGATGCGAAAAGCAGTCATGGACAGAAACTTCCTAG        13080
```

```
                                             End of the H-repeat
CGTCAGGCATTGCAGCGTGCGGGCTTTCATAATCTTGCAT TGGTTTTGATAAGATATTTC        13140
```

```
          Start of orf12
          M   N   L   Y   G   I   F   G   A   G   S   Y   G   R   E
TTTGGAGATGGGAAAATGAATTTGTATGGTATTTTTGGTGCTGGAAGTTATGGTAGAGAA        13200

    T   I   P   I   L   N   Q   Q   I   K   Q   E   C   G   S   D   Y   A   L   V
ACAATACCCATTCTAAATCAACAAATAAAGCAAGAATGTGGTTCTGACTATGCTCTGGTT        13260

    F   V   D   D   V   L   A   G   K   K   V   N   G   F   E   V   L   S   T   N
TTTGTGGATGATGTTTTGGCAGGAAAGAAAGTTAATGGTTTTGAAGTGCTTTCAACCAAC        13320

    C   F   L   K   A   P   Y   L   K   K   Y   F   N   V   A   I   A   N   D   K
TGCTTTCTAAAAGCCCCTTATTTAAAAAAGTATTTTAATGTTGCTATTGCTAATGATAAG        13380

    I   R   Q   R   V   S   E   S   I   L   L   H   G   V   E   P   I   T   I   K
ATACGACAGAGAGTGTCTGAGTCAATATTATTACACGGGGTTGAACCAATAACTATAAAA        13440

    H   P   N   S   V   V   Y   D   H   T   M   I   G   S   G   A   I   I   S   P
CATCCAAATAGCGTTGTTTATGATCATACTATGATAGGTAGTGGCGCTATTATTTCTCCC        13500

    F   V   T   I   S   T   N   T   H   I   G   R   F   F   H   A   N   I   Y   S
TTTGTTACAATATCTACTAATACTCATATAGGGAGGTTTTTTCATGCAAACATATACTCA        13560

    Y   V   A   H   D   C   Q   I   G   D   Y   V   T   F   A   P   G   A   K   C
TACGTTGCACATGATTGTCAAATAGGAGACTATGTTACATTTGCTCCTGGGGCTAAATGT        13620

    N   G   Y   V   V   I   E   D   N   A   Y   I   G   S   G   A   V   I   K   Q
AATGGATATGTTGTTATTGAAGACAATGCATATATAGGCTCGGGTGCAGTAATTAAGCAG        13680

    G   V   P   N   R   P   L   I   I   G   A   G   A   I   I   G   M   G   A   V
GGTGTTCCTAATCGCCCACTTATTATTGGCGCGGGAGCCATTATAGGTATGGGGGCTGTT        13740

    V   T   K   S   V   P   A   G   I   T   V   C   G   N   P   A   R   E   M   K
GTCACTAAAAGTGTTCCTGCCGGTATAACTGTGTGCGGAAATCCAGCAAGAGAAATGAAA        13800
```

```
                    End of orf12
    R   S   P   T   S   I   *
AGATCGCCAACATCTATT TAATGGGAATGCGAAAACACGTTCCAAATGGGACTAATGTTT        13860
```

Figure 8/11

09/423093

AAAATATATATAATTTCGCTAATTTACTAAATTATGGCTTCTTTTTAAGCTATCCTTTAC   13920

TTAGTTATTACTGATACAGCATGAAATTTATAATACTCTGATACATTTTTATACGTTATT   13980

CAAGCCGCATATCTAGCGGTAACCCCTGACAGGAGTAAACAATG   14024

Figure 8/12

66666

666676666

666666666666

```
GTTGACAAATACCGACCGTATAATGAATCAAACGTTCTGGATTGGTATTTATCCAGGCTT        60

GACTACAGAGCATTTAGATTATGTCGTAAGTAAGTTTGAAGAATTTTTTGGTTTAAATTT       120

            Start of abe
             M  L  D  V  N  K  K  I  L  M  T  G  A  T
CTAATTTTTAGGATAGGATGCTTGATGTGAATAAGAAAATCCTAATGACTGGCGCTACTA       180

 S  F  V  G  T  H  L  L  H  S  L  I  K  E  G  Y  S  I  I  A
GCTTTGTAGGTACCCATCTACTACATAGTCTCATAAAGGAAGGTTATAGTATTATTGCAT       240

 L  K  R  P  I  T  E  P  T  I  I  N  T  L  I  E  W  L  N  I
TAAAGCGTCCTATAACCGAGCCAACGATTATCAATACCTTGATTGAATGGTTGAATATAC       300

 Q  D  I  E  K  I  C  Q  S  S  M  N  I  H  A  I  V  H  I  A
AAGATATAGAAAAAATATGTCAATCATCTATGAATATTCATGCGATTGTCCATATTGCAA       360

 T  D  Y  G  R  N  R  T  P  I  S  E  Q  Y  K  C  N  V  L  L
CAGACTATGGTCGAAACAGAACCCCTATATCTGAACAATATAAATGTAATGTCCTATTAC       420

 P  T  R  L  L  E  L  M  P  A  L  K  T  K  F  F  I  S  T  D
CAACAAGACTGCTTGAGTTAATGCCAGCGCTTAAAACGAAATTCTTTATTTCTACTGACT       480

 S  F  F  G  K  Y  E  K  H  Y  G  Y  M  R  S  Y  M  A  S  K
CTTTTTTTGGGAAATATGAGAAGCACTATGGATATATGCGTTCTTACATGGCATCTAAAA       540

 R  H  F  V  E  L  S  K  I  Y  V  E  E  H  P  D  V  C  F  I
GACATTTTGTAGAACTATCAAAAATATACGTAGAGGAACATCCAGACGTTTGTTTTATAA       600

 N  L  R  L  E  H  V  Y  G  E  R  D  K  A  G  K  I  I  P  Y
ATTTACGTTTAGAACATGTTTACGGTGAGAGGGATAAAGCAGGTAAAATAATCCCGTATG       660

 V  I  K  K  M  K  N  N  E  D  I  D  C  T  I  A  R  Q  K  R
TTATCAAAAAAATGAAAAACAATGAAGATATTGATTGTACGATCGCCAGGCAGAAAAGAG       720

 D  F  I  Y  I  D  D  V  V  S  A  Y  L  K  I  L  K  E  G  F
ATTTTATTTATATAGACGATGTTGTTTCGGCCTATTTGAAAATTTTAAAGGAGGGTTTTA       780

 N  A  G  H  Y  D  V  E  V  G  T  G  K  S  I  E  L  K  E  V
ACGCTGGACACTATGATGTCGAGGTGGGGACTGGAAAATCGATAGAGCTAAAAGAAGTGT       840

 F  E  I  I  K  K  E  T  H  S  S  S  K  I  N  Y  G  A  V  A
TTGAGATAATAAAAAAAGAAACGCATAGTAGTAGTAAGATAAATTATGGTGCAGTTGCGA       900

 M  R  D  D  E  I  M  E  S  H  A  N  T  S  F  L  T  R  L  G
TGCGTGATGATGAGATTATGGAGTCACATGCAAATACCTCTTTCTTGACTCGATTAGGTT       960

                            End of abe   Start of wzx
                                                     M
 W  S  A  E  F  S  I  E  K  G  V  K  K  M  L  S  M  K  E  *
GGAGTGCCGAGTTTTCTATTGAGAAGGGTGTGAAAAAAATGTTGAGTATGAAAGAG TAAT     1020

    N  R  I  I  R  M  L  G  V  D  K  A  I  R  Y  V  I  F  G  K
GAATCGTATTATTAGAATGTTAGGTGTAGATAAAGCAATTCGTTATGTTATTTTTGGTAA     1080

    I  I  S  V  L  T  G  L  L  L  I  M  L  I  S  H  H  L  S  K
GATAATATCTGTATTAACGGGTTTACTGTTAATAATGTTAATATCACACCATTTATCTAA     1140

    D  A  Q  G  Y  Y  Y  T  F  N  S  V  V  A  L  Q  I  I  F  E
AGACGCACAGGGCTATTATTATACATTTAATTCAGTAGTGGCACTACAGATAATATTTGA     1200
```

Figure 9/1

```
      L  G  L  S  T  V  I  I  Q  F  A  S  H  E  M  S  A  L  K  Y
     ATTGGGGCTATCAACGGTAATCATTCAATTCGCTAGCCATGAAATGTCAGCGTTAAAATA      1260

      D  Y  S  E  R  D  I  I  G  E  S  K  N  K  Q  R  Y  L  S  L
     TGATTATTCTGAACGAGATATTATAGGTGAAAGTAAAAATAAGCAACGTTACCTATCGTT       1320

      F  R  L  A  I  K  W  Y  A  V  I  A  L  L  I  I  L  I  V  G
     ATTTCGGTTGGCAATAAAATGGTATGCAGTAATAGCTTTGCTAATAATATTAATAGTCGG      1380

      P  I  G  Y  V  F  F  T  Q  K  E  G  L  G  V  P  W  Q  G  A
     TCCCATCGGGTATGTTTTTTTTACGCAAAAAGAAGGCTTAGGTGTACCTTGGCAAGGGGC      1440

      W  L  L  L  T  I  V  T  A  F  N  I  F  L  V  S  V  L  S  V
     ATGGTTATTATTAACAATAGTTACAGCTTTTAATATTTTTCTTGTTTCTGTACTTTCTGT     1500

      A  E  G  S  G  L  I  T  D  V  N  K  M  R  M  Y  Q  S  L  L
     CGCTGAAGGGAGTGGGTTAATTACTGATGTGAATAAAATGAGAATGTATCAGTCGCTGTT      1560

      A  G  I  L  A  V  S  L  L  I  S  G  F  G  L  Y  A  T  S  A
     AGCTGGTATATTGGCAGTAAGCTTACTTATTAGTGGCTTTGGACTATATGCTACGTCTGC     1620

      I  A  I  S  G  T  I  I  F  S  I  F  S  Y  K  Y  F  K  K  I
     AATAGCTATTTCAGGGACTATCATATTCTCCATATTTTCATATAAGTATTTTAAAAAAAT    1680

      F  L  Q  S  L  K  H  K  N  K  Y  T  E  G  G  I  S  W  V  N
     TTTCCTGCAATCTTTAAAGCATAAAAATAAATATACTGAAGGTGGTATTTCATGGGTTAA      1740

      E  I  F  P  M  Q  W  R  I  A  L  S  W  M  S  G  Y  F  I  Y
     TGAAATATTTCCTATGCAATGGCGAATTGCTCTAAGTTGGATGTCAGGGTATTTTATTTA      1800

      F  V  M  T  P  I  A  F  K  Y  F  G  A  I  Y  A  G  Q  L  G
     TTTTGTTATGACCCCCATTGCATTCAAATATTTCGGGGCTATATATGCAGGGCAGTTAGG     1860

      M  S  L  T  L  C  N  M  V  M  A  T  G  L  A  W  I  S  T  K
     GATGTCTTTAACATTATGCAATATGGTAATGGCTACGGGCCTGGCTTGGATATCCACTAA      1920

      Y  P  K  W  G  V  M  V  S  N  K  Q  L  A  E  L  S  K  S  F
     ATATCCAAAATGGGGAGTAATGGTTTCCAACAAACAGCTTGCGGAACTGAGTAAATCGTT      1980

      K  S  A  V  M  Q  S  S  F  F  V  L  T  G  L  T  G  V  Y  I
     CAAAAGTGCAGTAATGCAATCATCCTTTTTTGTCTTGACAGGATTAACTGGTGTATACAT      2040

      S  L  W  L  L  K  L  S  G  S  N  I  G  E  R  F  L  G  L  Q
     TTCATTATGGTTATTGAAATTATCTGGTTCAAACATTGGCGAGCGGTTTTTGGGATTGCA      2100

      D  F  F  F  L  S  L  A  I  I  G  N  H  I  V  A  C  F  A  T
     GGATTTTTTCTTTTTATCTTTAGCAATTATTGGTAATCACATTGTAGCTTGCTTTGCAAC      2160

      Y  I  R  A  H  K  T  E  K  M  T  L  A  S  C  I  M  A  L  L
     CTATATAAGAGCGCATAAAACTGAAAAAATGACATTGGCATCATGTATAATGGCTCTCTT      2220

      T  I  T  T  M  L  F  V  A  Y  L  E  Y  S  R  F  Y  M  L  M
     GACTATAACTACAATGTTGTTTGTTGCATATTTAGAGTACTCGAGGTTCTACATGTTAAT      2280

      Y  A  A  L  T  W  L  Y  F  V  P  Q  T  Y  I  I  F  K  R  F
                                                    S  L  K  D
     GTATGCAGCACTAACGTGGTTATATTTTGTTCCTCAAACTTATATAATCTTTAAAAGATT      2340
```

Figure 9/2

```
          Start of wbaR   End of wzx
   K  S  S  Y  E  *
                 M  S  K  K  P  L  L  T  I  A  I  P  T  Y  N  R
CAAGAGTTCTTATGAGTAAAAAACCTCTTCTTACTATTGCTATTCCGACATATAACCGCT        2400

   S  S  C  L  A  R  L  L  D  S  I  I  Q  Q  E  N  Y  C  H  D
CTTCATGTTTGGCTCGTTTACTTGATAGTATAATTCAACAGGAGAACTATTGTCATGATG        2460

   E  L  E  V  I  V  C  D  N  A  S  T  D  E  T  A  R  I  A  K
AACTCGAGGTTATTGTTTGTGATAATGCTTCAACAGATGAAACAGCAAGAATAGCCAAGA        2520

   S  G  L  D  K  I  R  N  S  T  Y  H  L  N  E  E  N  L  G  M
GTGGCTTAGATAAAATAAGAAATAGTACTTATCATCTAAATGAAGAAAACTTAGGAATGG        2580

   D  G  N  F  Q  K  C  F  E  L  S  N  G  K  Y  L  W  M  I  G
ATGGTAACTTCCAGAAATGTTTTGAGTTATCAAATGGAAAATATCTTTGGATGATTGGCG        2640

   D  D  D  L  I  V  K  N  G  I  S  K  V  F  S  I  L  K  S  R
ATGATGATCTAATAGTCAAAAATGGTATTTCGAAGGTTTTTTCGATATTAAAGTCCCGGC        2700

   P  A  L  D  M  V  Y  V  N  S  A  A  K  T  E  L  N  Y  N  A
CTGCATTAGATATGGTGTATGTAAATTCAGCAGCAAAGACTGAGTTAAACTATAATGCTG        2760

   D  V  R  T  S  F  Y  T  N  D  V  D  F  I  S  D  V  K  V  M
ATGTGAGGACGTCATTCTACACAAATGATGTAGATTTTATTTCAGACGTGAAAGTTATGT        2820

   F  T  F  I  S  G  M  I  C  K  K  T  D  A  I  V  K  A  V  G
TCACGTTTATTTCTGGAATGATATGTAAGAAAACTGATGCAATTGTCAAAGCCGTTGGTA        2880

   I  F  S  P  Q  T  T  G  K  Y  L  M  H  L  T  W  Q  L  P  L
TTTTCAGTCCGCAAACTACTGGAAAATATCTTATGCATTTAACATGGCAATTGCCATTAC        2940

   L  K  Q  G  G  E  F  A  V  I  H  N  N  I  I  E  A  E  P  D
TTAAACAGGGTGGAGAGTTCGCAGTTATCCATAATAATATAATTGAGGCTGAGCCAGATA        3000

   N  S  G  G  Y  H  L  Y  K  V  F  S  N  N  L  A  T  I  F  D
ATTCAGGTGGATATCATTTATATAAGGTTTTTTCTAATAATCTTGCGACAATCTTTGATG        3060

   V  F  Y  P  R  E  H  R  V  S  K  R  V  R  A  S  A  C  L  F
TTTTTTATCCCAGAGAGCACCGTGTAAGTAAAAGAGTTCGCGCATCAGCATGTTTATTCT        3120

   L  L  N  F  I  G  D  E  D  K  T  K  N  F  A  T  N  N  Y  L
TACTTAACTTCATAGGCGATGAAGATAAAACCAAAAATTTTGCTACAAATAATTATTTAA        3180

   R  D  C  D  S  A  F  I  D  L  I  I  Y  K  Y  G  L  R  F  F
GAGATTGCGATAGTGCATTTATAGATTTAATTATATATAAATATGGGCTTAGGTTTTTCT        3240

   Y  L  Y  P  K  T  V  P  L  F  R  K  I  K  Y  I  I  K  T  V
ATCTATATCCTAAAACTGTGCCTTTATTTAGAAAAATAAAATATATTATAAAGACGGTTT        3300

          End of wbaR
   L  M  R  K  *
TAATGCGGAAATAAAAATTATTCAAGATGGTTTGCTGAAAACGACTTATAGGACTATCTA        3360

       Start of wbaL
   M  F  V  Y  S  L  R  L  K  L  N  L  I  I  S  L  L  S  K  V
ATGTTTGTCTATAGTTTAAGATTAAAATTAAATCTTATCATATCATTATTGAGTAAAGTT        3420

       R  R  K  S  K  A  K  F  L  V  L  L  S  G  Y  D  F  K  M  V
AGGCGGAAATCAAAAGCAAAGTTTCTTGTTCTGCTTAGCGGATATGATTTTAAAATGGTT        3480
```

Figure 9/3

```
       G   K   N   F   K   L   N   V   K   P   Y   S   A   K   N   N   T   S   S   K
      GGGAAGAATTTTAAATTGAATGTCAAACCTTACTCTGCAAAAAATAACACCTCTTCCAAA                    3540

       W   G   S   M   R   V   G   D   N   C   W   I   E   A   V   Y   N   Y   G   D
      TGGGGTAGTATGCGGGTTGGTGATAACTGCTGGATTGAAGCTGTATATAATTATGGTGAT                    3600

       E   K   F   E   P   Y   L   Y   I   G   D   R   I   C   L   S   D   N   V   H
      GAAAAATTTGAACCTTATTTGTACATAGGTGATCGTATATGTTTAAGTGATAATGTTCAT                    3660

       I   S   C   V   S   C   L   I   L   E   N   D   I   L   I   G   S   K   V   Y
      ATTTCTTGCGTATCATGTTTAATTTTAGAAAACGATATATTAATTGGTAGCAAAGTTTAT                    3720

       I   G   D   H   S   H   G   S   Y   K   V   C   S   P   K   I   E   P   P   A
      ATAGGCGATCATAGCCATGGCAGTTATAAAGTATGCAGTCCGAAAATAGAACCGCCAGCA                    3780

       N   K   P   L   G   D   I   A   P   I   K   I   G   N   C   C   W   I   G   D
      AATAAGCCATTAGGTGATATTGCTCCTATTAAAATAGGTAATTGCTGCTGGATTGGAGAT                    3840

       N   A   V   I   L   A   G   S   E   I   C   D   G   C   V   I   A   A   N   S
      AATGCAGTAATTCTGGCTGGTAGTGAAATTTGTGATGGCTGTGTAATCGCAGCTAATTCA                    3900

       V   V   K   D   L   K   V   D   K   P   C   L   I   G   G   V   P   A   K   V
      GTCGTCAAGGATTTAAAAGTCGATAAGCCATGTTTAATTGGTGGGGTTCCTGCTAAAGTA                    3960

              End of wbaL   Start of wbaQ
       I   K   V   F   *
                           M   N   V   F   I   S   I   C   I   P   S   Y   N   R   A
      ATAAAGGTATTTTAAAATGAATGTTTTTATCAGTATTTGTATACCGTCTTATAATAGAGC                    4020

       E   F   L   E   P   L   L   D   S   I   Y   N   Q   D   Y   C   L   K   N   N
      TGAGTTTTTAGAGCCACTACTGGATAGCATATATAATCAAGATTATTGTTTAAAGAATAA                    4080

       D   F   E   V   I   V   C   E   D   K   S   P   Q   R   D   E   I   N   S   I
      TGATTTTGAGGTCATTGTTTGTGAAGATAAATCTCCACAGAGAGATGAGATAAACTCTAT                    4140

       I   E   N   Y   K   A   K   N   N   K   Q   N   L   Y   V   N   F   N   E   D
      TATCGAAAACTATAAAGCAAAAAAATAATAAACAAAATCTTTATGTTAATTTCAATGAAGA                    4200

       N   L   G   Y   D   K   N   L   K   K   C   I   S   L   T   T   G   K   Y   C
      TAATTTAGGCTATGATAAGAATTTAAAAAAATGCATTAGTTTGACGACAGGTAAATATTG                    4260

       M   I   M   G   N   D   D   L   L   A   D   G   A   L   S   K   I   V   K   V
      CATGATCATGGGCAACGATGATCTATTAGCAGATGGAGCGTTATCAAAAATAGTGAAAGT                    4320

       L   K   A   N   P   E   I   V   L   A   T   R   A   Y   G   W   F   K   E   N
      TTTGAAGGCTAATCCTGAAATTGTATTGGCTACGCGAGCGTATGGTTGGTTTAAGGAAAA                    4380

       P   N   E   L   C   D   T   V   R   H   L   T   D   D   T   L   F   Q   P   G
      TCCGAATGAGTTATGTGATACTGTTCGTCATTTAACAGACGATACTTTATTTCAGCCGGG                    4440

       A   D   A   I   K   F   F   F   R   R   V   G   V   I   S   G   F   I   V   N
      GGCTGATGCCATTAAATTTTTCTTCCGTAGAGTTGGAGTTATTTCAGGCTTTATTGTCAA                    4500

       A   E   K   A   K   K   L   S   S   D   L   F   D   G   R   L   Y   Y   Q   M
      TGCTGAAAAAGCAAAAAAACTATCGAGTGATTTATTTGATGGGCGTTTATATTATCAAAT                    4560

       Y   L   A   G   M   L   M   A   E   G   Q   G   Y   Y   F   S   D   V   M   T
      GTACCTTGCTGGTATGCTAATGGCTGAAGGTCAGGGATACTATTTTAGCGACGTGATGAC                    4620
```

Figure 9/4

```
   L   S   R   D   T   E   A   P   D   F   G   N   A   G   T   E   K   G   V   F
ATTGTCGAGGGATACAGAGGCTCCTGACTTTGGTAACGCTGGAACTGAAAAAGGAGTTTT        4680

   T   P   G   G   Y   K   P   E   G   R   I   H   M   V   E   G   L   L   L   I
CACCCCGGGGGGGTATAAACCAGAGGGCCGTATACATATGGTTGAAGGCTTGTTGCTAAT        4740

   A   K   Y   I   E   D   T   T   K   I   D   G   V   Y   A   G   I   R   K   D
TGCAAAATATATAGAAGATACAACAAAAATTGATGGCGTTTATGCTGGAATTAGAAAAGA        4800

   L   A   N   Y   F   Y   P   Y   I   R   D   Q   L   D   L   P   L   Y   T   Y
CTTAGCGAACTATTTTTATCCTTATATTCGAGATCAACTCGACTTGCCTCTTTATACTTA        4860

   I   K   M   I   N   K   F   R   K   M   G   F   S   N   E   K   L   F   Y   V
TATTAAAATGATAAATAAATTTCGGAAAATGGGATTTTCAAATGAAAGCTTTTCTATGT        4920

   H   A   F   L   G   Y   V   L   K   R   R   G   Y   D   A   L   I   K   Y   I
GCATGCCTTTTTAGGGTATGTACTAAAACGGAGGGGCTATGATGCTTTAATTAAATACAT        4980

                                 End of wbaQ
   R   S   K   K   G   G   T   P   R   L   G   I   *
TCGTAGCAAAAAAGGCGGTACTCCGCGTCTTGGTATT TAACCTCCACTTTCAAAAAATGT        5040

TATGAATATACTTCTTGCTGCGATATTAGGCGTTAACTTATTTTCTCCATATATTAGTTC        5100

        Start of wzy
                 M   L   P   F   P   P   G   A   I   L   R   D   V   L   N   V
GTGGATGGTGGGTATGCTGCCATTTCCACCAGGAGCAATCCTAAGGGATGTACTCAATGT        5160

   F   F   V   A   L   V   L   V   R   F   V   I   D   R   K   K   T   Y   F   P
ATTTTTTGTGGCGTTAGTGCTAGTTCGATTTGTCATTGATAGGAAAAAAACTTATTTCCC        5220

   L   V   F   T   I   F   S   W   S   A   V   I   L   W   V   I   A   L   T   I
GTTGGTTTTTACTATTTTTTCATGGTCGGCGGTAATACTATGGGTAATAGCGTTAACTAT        5280

   F   S   P   D   K   I   Q   A   I   M   G   G   R   S   Y   I   L   F   P   A
ATTCTCACCGGATAAAATTCAAGCAATTATGGGGGGGCGGAGTTATATTTTATTCCCGGC        5340

   V   F   I   A   L   V   I   L   K   V   S   Y   P   Q   S   L   N   I   E   K
AGTTTTCATAGCATTAGTGATTTTAAAAGTATCATACCCGCAATCCTTAAATATTGAAAA        5400

   I   V   C   Y   I   I   F   L   M   F   M   V   A   T   I   S   I   I   D   V
AATAGTTTGCTACATAATTTTTCTAATGTTTATGGTTGCGACAATATCTATTATTGATGT        5460

   L   M   N   G   E   F   I   K   L   L   G   Y   D   E   H   Y   A   G   E   Q
ACTAATGAATGGAGAGTTCATTAAATTGCTCGGATATGATGAGCATTATGCAGGAGAACA        5520

   L   N   L   I   N   S   Y   D   G   M   V   R   A   T   G   G   F   S   D   A
ATTAAACTTAATTAATAGCTATGATGGGATGGTCCGGGCTACAGGCGGTTTTAGTGATGC        5580

   L   N   F   G   Y   M   L   T   L   G   V   L   L   C   M   E   C   F   S   Q
TCTCAATTTTGGATATATGCTCACATTAGGTGTTTTGTTATGTATGGAGTGTTTTCCCA        5640

   G   Y   K   R   L   L   M   L   I   I   S   F   V   L   F   I   A   I   C   M
AGGATATAAAAGATTATTGATGCTTATTATTAGTTTTGTGCTATTTATAGCGATCTGCAT        5700

   S   L   T   R   G   A   I   L   V   A   A   L   I   Y   A   L   Y   I   I   S
GAGTCTTACTAGAGGAGCAATACTTGTTGCTGCGCTTATTTACGCACTTTATATAATTTC        5760

   N   R   K   M   L   F   C   G   I   T   L   F   V   I   I   I   P   V   L   A
AAATCGGAAGATGCTTTTTTTGTGGAATAACTTTATTTGTAATAATTATACCCGTTTTAGC        5820
```

Figure 9/5

09/423093

```
      I   S   T   N   I   F   D   N   Y   T   E   Y   L   I   G   R   F   T   D   S
AATTTCTACTAATATTTTTGACAACTATACAGAAATTTTGATCGGCAGGTTTACAGATTC          5880

      S   Q   A   S   R   G   S   T   Q   G   R   I   D   M   A   I   N   S   L   N
GTCTCAGGCATCGCGTGGATCTACACAGGGGCGGATAGATATGGCAATTAATTCATTAAA          5940

      F   L   S   E   H   P   S   G   I   G   L   G   T   Q   G   S   G   N   M   L
CTTCCTGTCAGAACATCCATCAGGTATAGGTCTGGGTACTCAAGGTTCAGGAAACATGCT          6000

      S   V   K   D   N   R   L   N   T   D   N   Y   F   F   W   I   A   L   E   T
TTCGGTAAAAGATAATAGGTTAAATACGGATAATTATTTTTTCTGGATCGCCCTTGAGAC          6060

      G   I   I   G   L   I   I   N   I   I   Y   L   A   S   Q   F   Y   S   S   T
TGGTATTATTGGCTTAATCATAAATATTATTTATCTGGCAAGTCAATTTTATTCTTCAAC          6120

      L   L   N   R   I   Y   G   S   H   C   S   N   M   H   Y   R   L   Y   F   L
TTTACTAAATAGAATATATGGCAGTCATTGTAGCAATATGCACTATAGATTATATTTTCT          6180

      F   G   S   I   Y   F   I   S   A   A   L   S   S   A   P   S   S   S   T   F
CTTTGGAAGTATATATTTTATAAGTGCAGCGTTAAGTTCAGCACCTTCGTCATCAACTTT          6240

      S   I   Y   Y   W   T   V   L   A   L   I   P   F   L   K   L   T   N   R   R
TTCTATATATTATTGGACAGTTTTAGCTTTGATTCCATTTTTAAAATTAACAAATAGACG          6300
```

End of *wzy* Start of *wbaW*

```
      C   T   R   *   M   N   N   K   K   V   L   M   D   I   S   W   S   N   K   G
GTGCACGCGA TAATGAATAATAAAAAGGTTTTGATGGATATTAGTTGGTCTAATAAAGGG          6360

      G   I   G   R   F   T   D   E   I   S   K   L   L   C   D   I   S   K   E   E
GGGATTGGACGTTTTACTGATGAAATTTCTAAACTACTATGTGATATATCTAAGGAGGAA          6420

      L   Y   R   K   C   A   S   P   L   A   P   L   G   L   A   V   N   I   F   L
CTATATAGAAAATGTGCTTCTCCGCTGGCCCCATTAGGTTTAGCAGTCAATATTTTTCTG          6480

      R   K   K   T   D   V   V   F   L   P   G   Y   I   P   P   L   F   C   S   K
CGAAAGAAAACTGATGTGGTTTTTCTTCCTGGCTATATTCCACCACTTTTTTGTTCGAAA          6540

      K   F   I   I   T   I   H   D   L   N   H   L   D   L   N   D   N   S   S   L
AAGTTCATAATAACAATACATGATCTAAATCATCTGGATTTAAATGATAATTCCTCTCTT          6600

      F   K   R   L   F   Y   N   F   I   I   K   R   G   C   R   K   A   Y   K   I
TTTAAGAGGTTATTTTATAATTTTATAATAAAGCGCGGTTGTAGAAAAGCATATAAAATA          6660

      F   T   V   S   N   F   S   K   E   R   I   V   A   W   S   G   V   N   P   N
TTTACAGTTTCGAATTTTTCAAAAGAAAGAATAGTAGCATGGTCAGGTGTAAACCCTAAT          6720

      K   I   V   T   V   Y   N   G   V   S   S   L   F   N   A   D   V   K   P   L
AAAATAGTCACGGTATATAATGGGGTATCTAGTCTATTTAATGCCGATGTAAAACCATTG          6780

      N   L   G   Y   K   Y   L   L   C   V   G   N   R   K   T   H   K   N   E   K
AATTTAGGCTATAAATATTTGCTATGTGTAGGAAACAGAAAAACTCATAAGAATGAGAAG          6840

      C   V   I   S   A   F   A   K   A   D   I   D   P   S   I   K   L   V   F   T
TGTGTTATATCTGCCTTTGCCAAAGCAGATATTGATCCATCAATAAAACTCGTTTTTACT          6900

      G   N   P   C   N   D   L   E   K   L   I   I   Q   H   G   L   S   E   R   V
GGTAATCCTTGTAATGATTTAGAAAAACTAATAATACAACATGGTTTAAGTGAACGTGTA          6960
```

Figure 9/6

```
     K  F  F  G  F  V  S  E  K  D  L  P  S  L  Y  K  G  S  L  G
     AAGTTCTTTGGGTTCGTGTCTGAAAAAGATTTACCATCGTTATATAAGGGCTCGTTAGGA    7020

     L  V  F  P  S  L  Y  E  G  F  G  L  P  V  V  E  G  M  A  C
     TTAGTTTTCCCTTCTTTATATGAAGGTTTTGGATTACCTGTAGTGGAGGGCATGGCCTGT    7080

     G  I  P  V  L  T  S  L  T  S  S  L  P  E  V  A  G  D  A  A
     GGTATTCCTGTATTAACTTCTCTAACTTCATCATTGCCAGAGGTGGCTGGAGATGCAGCG    7140

     I  L  V  D  P  L  S  E  D  A  I  T  K  G  I  S  R  L  I  N
     ATTCTTGTCGACCCTCTTTCGGAAGATGCTATTACTAAAGGAATTTCGAGGTTAATTAAT    7200

     D  S  E  L  R  K  H  L  I  Q  K  G  L  L  R  A  K  R  F  N
     GATTCTGAACTTCGTAAGCATTTAATCCAAAAGGGGCTTTTGCGGGCAAAGAGGTTCAAT    7260

                                                 Start of wbaZ
     W  Q  N  V  V  S  E  I  E  M  V  L  T  E  A  C  D  G  N  K
                                                         M  E  I  N
     TGGCAAAACGTGGTTAGTGAGATTGAAATGGTACTGACAGAGGCATGTGATGGAAATAAA    7320

     End of wbaW
     *
     E  I  K  I  S  L  V  H  E  W  L  L  S  Y  A  G  S  E  Q  V
     TGAAATAAAAATATCTCTCGTTCATGAGTGGTTATTAAGTTATGCAGGCTCCGAACAGGT    7380

     S  S  A  I  L  H  V  F  P  E  A  K  L  Y  S  V  V  D  F  L
     ATCATCTGCCATCCTGCATGTTTTTCCTGAAGCGAAGTTATATTCGGTGGTTGATTTTCT    7440

     T  D  E  Q  R  R  H  F  L  G  K  Y  A  T  T  T  F  I  Q  N
     AACGGATGAACAAAGAAGACATTTTCTGGGGAAATATGCGACTACCACATTTATTCAAAA    7500

     L  P  K  A  K  K  F  Y  Q  K  Y  L  P  L  M  P  L  A  I  E
     TTTACCTAAAGCTAAAAAATTTTACCAGAAATATTTACCACTAATGCCACTGGCTATTGA    7560

     Q  L  D  L  S  D  A  N  I  I  I  S  S  A  H  S  V  A  K  G
     ACAACTTGATTTATCAGATGCTAATATCATCATTAGTAGCGCCCATTCCGTTGCAAAAGG    7620

     V  I  S  G  P  D  Q  L  H  I  S  Y  V  H  S  P  I  R  Y  A
     TGTTATTTCCGGACCAGATCAGCTTCACATTAGCTATGTTCATTCTCCTATTCGATATGC    7680

     W  D  L  Q  H  Q  Y  L  N  E  S  N  L  N  K  G  I  K  G  W
     GTGGGATTTACAGCATCAGTACCTTAATGAGTCTAACCTGAATAAAGGAATTAAAGGTTG    7740

     L  A  K  W  L  L  H  K  I  R  I  W  D  S  R  T  A  N  G  V
     GTTAGCAAAATGGCTTCTTCACAAAATACGAATTTGGGATTCTCGAACCGCAAATGGGGT    7800

     D  H  F  I  A  N  S  Q  Y  I  A  R  R  I  K  K  V  Y  R  R
     TGATCATTTTATAGCTAATTCTCAATATATCGCGCGTAGAATTAAAAAAGTATACAGACG    7860

     E  A  S  V  I  Y  P  P  V  D  V  D  N  F  E  V  K  N  E  K
     TGAGGCTTCAGTTATATATCCGCCTGTAGATGTGGATAATTTTGAAGTAAAAAATGAAAA    7920

     Q  D  Y  Y  F  T  A  S  R  M  V  P  Y  K  R  I  D  L  I  V
     GCAAGACTATTATTTCACAGCATCCCGTATGGTACCCTACAAACGTATTGATCTTATTGT    7980

     E  A  F  S  K  M  P  E  K  K  L  V  V  I  G  D  G  P  E  M
     CGAAGCCTTTAGTAAAATGCCGGAAAAGAAATTAGTAGTTATTGGTGATGGACCGGAGAT    8040

     K  K  I  K  S  K  A  T  D  N  I  K  L  L  G  Y  Q  S  F  P
     GAAAAAAATAAAGAGCAAGGCTACAGACAATATAAAATTGCTCGGTTATCAATCTTTTCC    8100
```

Figure 9/7

```
          V  L  K  E  Y  M  Q  S  A  R  A  F  V  F  A  A  E  E  D  F
        TGTTTTAAAAGAGTATATGCAGAGCGCCAGGGCGTTTGTTTTTGCAGCGGAAGAGGACTT        8160

          G  I  I  P  V  E  A  Q  A  C  G  T  P  V  I  A  F  G  K  G
        TGGAATAATACCTGTCGAAGCTCAAGCTTGCGGTACCCCTGTTATTGCCTTTGGGAAGGG        8220

          G  A  L  E  T  V  R  P  L  G  V  E  E  P  T  G  I  F  F  K
        TGGGGCCTTAGAAACCGTTCGCCCACTAGGTGTAGAGGAACCGACTGGCATTTTCTTCAA        8280

          E  Q  N  I  A  S  L  H  E  A  V  S  E  F  E  K  N  A  S  F
        GGAACAGAATATTGCTTCTTTGCATGAAGCTGTTAGTGAATTTGAAAAAAATGCATCATT        8340

          F  T  S  Q  A  C  R  K  N  A  E  K  F  S  R  S  R  F  E  Q
        TTTTACATCTCAGGCTTGTAGAAAAAATGCAGAAAAATTTTCTCGATCAAGATTTGAACA        8400

          E  F  K  N  F  V  N  E  K  W  N  L  F  K  T  E  Q  I  I  K
        AGAATTTAAGAACTTTGTTAATGAAAGTGGAATCTTTTCAAAACAGAACAGATTATTAA  ·  8460

        End of wbaZ Start of manC
                          M  S  K  L  I  P  V  I  M  A  G  G  I
          R  ·*
        ACGTTAATTATGGTTTATTGAATGTCTAAATTAATACCAGTAATAATGGCCGGTGGGATT        8520

          G  S  R  L  W  P  L  S  R  E  E  H  P  K  Q  F  L  S  V  D
        GGTAGCCGTTTGTGGCCACTTTCACGTGAAGAGCATCCGAAACAGTTTTTAAGCGTAGAT        8580

          G  E  L  S  M  L  Q  N  T  I  K  R  L  T  P  L  L  A  G  E
        GGTGAATTATCTATGCTGCAAAACACCATTAAAAGATTGACTCCTCTTTTGGCTGGAGAA        8640

          P  L  V  I  C  N  D  S  H  R  F  L  V  A  E  Q  L  R  A  I
        CCTTTAGTCATTTGTAATGATAGTCACCGCTTCCTTGTCGCTGAACAACTTCGAGCTATA        8700

          N  K  L  A  N  N  I  I  L  E  P  V  G  R  N  T  A  P  A  I
        AATAAACTAGCAAATAACATCATATTAGAGCCAGTGGGGCGTAATACAGCCCCAGCTATA        8760

          A  L  A  A  F  C  S  L  Q  N  V  V  D  E  D  P  L  L  L  V
        GCGCTGGCCGCTTTTTGTTCACTTCAGAATGTCGTCGATGAAGACCCGCTTTTGCTTGTC        8820

          L  A  A  D  H  V  I  R  D  E  K  V  F  L  K  A  I  N  H  A
        CTTGCTGCGGATCATGTCATCCGCGATGAGAAAGTGTTTCTTAAAGCTATCAATCACGCT        8880

          E  F  F  A  T  Q  G  K  L  V  T  F  G  I  V  P  T  Q  A  E
        GAATTTTTTGCAACACAAGGTAAGCTAGTAACGTTTGGTATTGTACCCACACAGGCCGAA        8940

          T  G  Y  G  Y  I  C  R  G  E  A  I  G  E  D  A  F  S  V  A
        ACTGGCTACGGTTATATTTGTAGAGGTGAAGCAATCGGGGAAGATGCTTTTTCTGTAGCC        9000

          E  F  V  E  K  P  D  F  D  T  A  R  H  Y  V  E  S  E  K  Y
        GAATTTGTAGAGAAGCCTGATTTCGATACAGCGCGTCATTATGTAGAATCAGAGAAATAT        9060

          Y  W  N  S  G  M  F  L  F  R  A  S  S  Y  L  Q  E  L  K  D
        TATTGGAACAGCGGTATGTTCCTATTTCGTGCAAGTAGTTACTTACAAGAATTAAAGGAT        9120

          L  S  P  D  I  Y  Q  A  C  E  N  A  V  G  S  I  N  P  D  L
        CTGTCCCCCGATATTTACCAAGCATGTGAAAATGCGGTAGGGAGTATTAATCCTGATCTT        9180

          D  F  I  R  I  D  K  E  A  F  A  M  C  P  S  D  S  I  D  Y
        GATTTTATCCGTATTGATAAAGAAGCATTCGCAATGTGCCCTAGTGATTCTATCGATTAT        9240
```

Figure 9/8

```
      A   V   M   E   H   T   R   H   A   V   V   V   P   M   N   A   G   W   S   D
GCGGTAATGGAACATACTAGGCATGCAGTTGTCGTACCGATGAATGCCGGCTGGTCAGAT                      9300

      V   G   S   W   S   S   L   W   D   I   S   K   K   D   P   Q   R   N   V   L
GTGGGGTCATGGTCTTCACTGTGGGATATTTCTAAGAAAGATCCACAACGTAATGTATTA                      9360

      H   G   D   I   F   A   Y   N   S   K   D   N   Y   I   Y   S   E   K   S   F
CATGGCGATATTTTTGCATATAATAGTAAAGATAATTATATCTATTCTGAAAAATCGTTT                      9420

      I   S   T   I   G   V   N   N   L   V   I   V   Q   T   A   D   A   L   L   V
ATTAGTACAATCGGAGTAAATAATTTAGTTATCGTGCAGACAGCAGATGCATTATTAGTA                      9480

      S   D   K   D   S   V   Q   D   V   K   K   V   V   D   Y   L   K   A   N   N
TCTGATAAAGATTCAGTCCAGGATGTTAAAAAAGTTGTTGATTATTTAAAAGCTAATAAT                      9540

      R   N   E   H   K   K   H   L   E   V   F   R   P   W   G   K   F   S   V   I
AGAAACGAACATAAAAAACATTTAGAGGTTTTCCGACCGTGGGGAAAATTTAGCGTAATT                      9600

      H   S   G   D   N   Y   L   V   K   R   I   T   V   K   P   G   A   K   F   A
CATAGTGGCGATAATTATTTAGTTAAAAGAATAACTGTTAAACCAGGCGCGAAGTTTGCT                      9660

      A   Q   M   H   L   H   R   A   E   H   W   I   V   V   S   G   T   A   C   I
GCTCAGATGCATCTCCATCGTGCTGAGCATTGGATAGTGGTATCTGGTACTGCTTGTATT                      9720

      T   K   G   E   E   I   F   T   I   S   E   N   E   S   T   F   I   P   A   N
ACTAAGGGGGAAGAAATTTTTACAATTTCGGAGAATGAATCAACATTTATACCTGCTAAT                      9780

      T   V   H   T   L   K   N   P   A   T   I   P   L   E   L   I   E   I   Q   S
ACAGTTCATACGTTAAAAAAACCCCGCGACTATTCCATTAGAACTAATAGAAATTCAATCT                     9840

      G   T   Y   L   A   E   D   D   I   I   R   L   E   K   H   S   G   Y   L   E
GGCACCTATCTTGCGGAGGATGATATTATTCGCCTGGAGAAACATTCTGGATATCTGGAG                      9900
```

End of manC Start of manB

```
*
                  M   K   N   I   Y   N   T   Y   D   V   I   N   K   S   G   I   N
TAATGAATTGATGAAAAATATATATAATACTTACGATGTTATCAACAAATCTGGAATTAA                      9960

      F   G   T   S   G   A   R   G   L   V   T   D   F   T   P   E   V   C   A   R
TTTTGGAACCAGTGGTGCCCGCGGCCTTGTTACCGATTTTACACCCGAAGTTTGCGCACG                     10020

      F   T   I   S   F   L   T   V   M   Q   Q   R   F   S   F   T   T   V   A   L
ATTTACCATTTCCTTTTTGACAGTAATGCAGCAAAGATTCTCATTTACAACGGTTGCGCT                     10080

      A   I   D   N   R   P   S   S   Y   A   M   A   Q   A   C   A   A   A   L   Q
CGCAATTGATAATCGTCCAAGCAGTTACGCGATGGCTCAAGCTTGTGCCGCTGCTTTGCA                     10140

      E   K   G   I   K   T   V   Y   Y   G   V   I   P   T   P   A   L   A   H   Q
AGAAAAAGGAATTAAAACCGTTTACTATGGCGTAATTCCAACACCTGCTTTAGCTCATCA                     10200

      S   I   S   D   K   V   P   A   I   M   V   T   G   S   H   I   P   F   D   R
ATCAATTTCCGATAAAGTACCTGCAATCATGGTTACTGGCAGTCATATCCCTTTTGACCG                     10260

      N   G   L   K   F   Y   R   P   D   G   E   I   T   K   D   D   E   N   A   I
TAATGGCCTGAAATTTTATAGACCAGATGGTGAAATTACTAAAGATGATGAGAATGCTAT                     10320

      I   H   V   D   A   S   F   M   Q   P   K   L   E   Q   L   T   I   S   T   I
TATTCATGTTGATGCCTCATTTATGCAGCCTAAGCTTGAACAATTGACAATTTCCACAAT                     10380
```

Figure 9/9

```
         A   A   R   N   Y   I   L   R   Y   T   S   L   F   P   M   P   F   L   K   N
       CGCTGCTAGAAATTATATTCTACGATATACCTCATTATTTCCAATGCCATTCTTGAAAAA        10440

           K   R   I   G   I   Y   E   H   S   S   A   G   R   D   L   Y   K   T   L   F
       TAAGCGCATTGGAATTTATGAGCATTCTAGTGCGGGTCGTGATCTCTATAAGACGTTATT        10500

           K   M   L   G   A   T   V   V   S   L   A   R   S   D   E   F   V   P   I   D
       CAAAATGTTGGGTGCTACAGTTGTTAGTTTAGCAAGGAGCGACGAATTTGTTCCTATTGA        10560

           T   E   A   V   S   E   D   D   R   N   K   A   I   T   W   A   K   K   Y   Q
       TACTGAAGCTGTAAGTGAAGATGATAGAAATAAAGCAATCACATGGGCAAAAAAATATCA        10620

           L   D   A   I   F   S   T   D   G   D   G   D   R   P   L   I   A   D   E   Y
       GTTAGATGCTATATTTTCAACTGATGGTGATGGAGATCGCCCTCTGATAGCTGACGAATA        10680

           G   N   W   L   R   G   D   I   L   G   L   L   C   S   L   E   L   A   A   D
       TGGAAATTGGTTAAGAGGAGATATATTAGGCCTTCTGTGCTCTCTCGAATTAGCTGCTGA        10740

           A   V   A   I   P   V   S   C   N   S   T   I   S   S   G   N   F   F   K   H
       TGCAGTCGCTATTCCTGTAAGCTGCAACAGTACAATCTCATCTGGTAACTTTTTTAAACA        10800

           V   E   R   T   K   I   G   S   P   Y   V   I   A   A   F   A   K   L   S   A
       TGTGGAACGAACAAAGATTGGTTCACCCTATGTGATTGCAGCATTTGCTAAATTATCTGC        10860

           N   Y   N   C   I   A   G   F   E   A   N   G   G   F   L   L   G   S   D   V
       AAACTATAATTGTATAGCTGGTTTTGAAGCGAATGGTGGCTTTCTGCTAGGTAGCGATGT        10920

           Y   I   N   Q   R   L   L   K   A   L   P   T   R   D   A   L   L   P   A   I
       TTATATTAATCAGCGTTTACTTAAGGCATTACCAACACGTGATGCTTTATTACCTGCCAT        10980

           M   L   L   F   G   S   K   D   K   S   I   S   E   L   V   K   K   L   P   A
       TATGCTTCTGTTTGGTAGCAAGGACAAAAGTATTAGTGAGCTTGTTAAAAAACTTCCTGC        11040

           R   Y   T   Y   S   N   R   L   Q   D   I   S   V   K   T   S   M   S   L   I
       TCGCTATACCTATTCAAACAGATTACAGGATATAAGTGTTAAAACAAGTATGTCTTTAAT        11100

           N   L   G   L   T   D   Q   E   D   F   L   Q   Y   I   G   F   N   K   H   H
       AAATCTTGGTCTGACAGATCAAGAGGATTTTTTGCAGTATATTGGTTTTAATAAACATCA        11160

           I   L   H   S   D   V   T   D   G   F   R   I   T   I   D   N   N   N   I   I
       TATATTACATTCTGATGTTACTGATGGCTTTAGAATCACTATCGATAACAACAATATTAT        11220

           H   L   R   P   S   G   N   A   P   E   L   R   C   Y   A   E   A   D   S   Q
       TCATTTACGACCTTCAGGCAATGCCCCTGAGTTGCGTTGCTATGCGGAGGCTGACTCGCA        11280

           E   D   A   C   N   I   V   E   T   V   L   S   N   I   K   S   K   L   G   R
       AGAGGATGCATGTAATATTGTTGAAACTGTTCTCTCTAATATCAAAAGCAAACTGGGTAG        11340
```

End of manB
```
           A   *
       AGCT TAATGCTGTTGATAATAGAGCGTTTCTTTCCAGTAATACTTTGTCTGGTTATCTGG        11400
```

Start of wbaP
```
                                                  M   D   R   F   D   N   K   Y   N   P   N   L
       TACCCAAGTTGAGGGTGAGAATTAAAATGGATCGTTTTGATAATAAGTATAACCCAAATTT        11460

           C   K   I   L   L   A   I   S   D   L   L   F   F   N   V   A   L   W   A   S
       ATGCAAAATATTATTGGCTATATCAGATTTACTGTTTTTTAATGTAGCCTTATGGGCATC        11520
```

Figure 9/10

```
      L   G   V   V   Y   L   I   F   D   E   V   Q   R   F   V   P   Q   E   Q   L
     GTTAGGAGTTGTATATTTAATCTTTGATGAAGTTCAGCGATTTGTACCACAAGAGCAATT          11580

      D   N   R   F   I   S   H   F   I   L   S   I   V   C   V   G   W   F   W   V
     AGATAATCGATTTATATCACATTTTATTCTATCTATAGTATGCGTTGGATGGTTTTGGGT          11640

      R   L   R   H   Y   T   Y   R   K   P   F   W   Y   E   L   K   E   V   I   R
     TCGACTGCGTCACTATACATATCGAAAGCCATTCTGGTATGAGTTGAAAGAGGTTATTCG          11700

      T   I   V   I   F   A   V   F   D   L   A   L   I   A   F   T   K   W   Q   F
     TACTATCGTTATTTTTGCTGTGTTTGATTTGGCTTTAATTGCGTTTACAAAATGGCAGTT          11760

      S   R   Y   V   W   V   F   C   W   T   F   A   I   I   L   V   P   F   F   R
     TTCACGCTATGTCTGGGTGTTTTGTTGGACTTTTGCCATAATCCTGGTGCCTTTTTTTCG          11820

      A   L   T   K   H   L   L   N   K   L   G   I   W   K   K   K   T   I   I   L
     CGCACTTACAAAGCATTTATTGAACAAGCTAGGTATCTGGAAGAAAAAAACTATCATCCT          11880

      G   S   G   Q   N   A   R   G   A   Y   S   A   L   Q   S   E   E   M   M   G
     TGGGAGCGGACAGAATGCTCGTGGTGCATATTCTGCGCTGCAAAGTGAGGAGATGATGGG          11940

      F   D   V   I   A   F   F   D   T   D   A   S   D   A   E   I   N   M   L   P
     GTTTGATGTTATCGCTTTTTTTTGATACGGATGCGTCAGATGCTGAAATAAATATGTTGCC          12000

      V   I   K   D   T   E   T   I   W   D   L   N   R   T   G   D   V   H   Y   I
     GGTGATAAAGGACACTGAGACTATTTGGGATTTAAATCGTACAGGTGATGTCCATTATAT          12060

      L   A   Y   E   Y   T   E   L   E   K   T   H   F   W   L   R   E   L   S   K
     CCTTGCTTATGAATACACCGAGTTGGAGAAAACACATTTTTGGCTACGTGAACTTTCAAA          12120

      H   H   C   R   S   V   T   V   V   P   S   F   R   G   L   P   L   Y   N   T
     ACATCATTGTCGTTCTGTTACTGTCGTCCCCTCGTTTAGAGGATTGCCATTATATAATAC          12180

      D   M   S   F   I   F   S   H   E   V   M   L   L   R   I   Q   N   N   L   A
     TGATATGTCTTTTATCTTTAGCCATGAAGTTATGTTATTAAGGATACAAAATAACTTGGC          12240

      K   R   S   S   R   F   L   K   R   T   F   D   I   V   C   S   I   M   I   L
     TAAAAGGTCGTCCCGTTTTCTCAAACGGACATTTGATATTGTTTGTTCAATAATGATTCT          12300

      I   I   A   S   P   L   M   I   Y   L   W   Y   K   V   T   R   D   G   G   P
     TATAATTGCATCACCACTTATGATTTATCTGTGGTATAAAGTTACTCGAGATGGTGGTCC          12360

      A   I   Y   G   H   Q   R   V   G   R   H   G   K   L   F   P   C   Y   K   F
     GGCTATTTATGGTCACCAGCGAGTAGGTCGGCATGGAAAACTTTTTCCATGCTACAAATT          12420

      R   S   M   V   M   N   S
     TCGTTCTATGGTTATGAATTC    12441
```

Figure 9/11

```
GAATTCGGGAGGCGCAATGAAAGTCAGCTTTTTTCTGCTGAAATTTCCACTCTCATCGGA        60
AACCTTTGTGCTGAATCAGATTACTGCGTTTATTGATATGGGCCATGAGGTGGAGATTGT       120
CGCGTTACAAAAAGGCGATACCCAACATACTCACGCCGCCTGGGAGAAGTATGGCCTGGC       180
GGCGAAAACCCGCTGGTTACAGGATGAGCCCCAGGGACGGCTGGCGAAACTGCGCTACCG       240
GGCATGTAAAACGCTGCCGGGGCTGCATCGGGCGGCGACCTGGAAAGCGCTCAATTTTAC       300
CCGCTATGGCGATGAATCACGCAATTTGATCCTTTCCGCGATTTGCGCGCAGGTGAGCCA       360
GCCTTTTGTGGCGGATGTGTTTATCGCACACTTTGGTCCGGCGGGCGTGACGGCGGCCAA       420
ACTACGCGAACTGGGCGTGCTTCGCGGCAAAATCGCGACTATTTTCCACGGGATTGATAT       480
CTCTAGTCGTGAGGTGCTCAGTCATTACACGCCGGAGTATCAGCAGTTGTTTCGTCGTGG       540
CGATCTGATGCTGCCCATCAGCGATCTGTGGGCCGGTCGCCTGAAAAGTATGGGCTGTCC       600
GCCGGAAAAGATTGCCGTTTCGCGCATGGGCGTCGACATGACGCGTTTTACCCATCGTTC       660
GGTGAAAGCGCCAGGGATGCCGCTGGAGATGATTTCCGTCGCGCGCCTGACAGAAAAAAA       720
AGGCCTGCATGTGGCGATTGAAGCCTGTCGGCAACTGAAAGCACAGGGCGTGGCGTTTCG       780
CTACCGCATTCTGGGGATTGGCCCGTGGGAACGTCGGCTGCGCACGCTCATCGAGCAGTA       840
TCAGCTAGAGGATGTCATTGAGATGCCGGGGGTTTAAAACCGAGCCATGAAGTGAAGGCGAT       900
GCTGGATGACGCCGATGTTTTTTTGCTGCCGTCGATTACCGGTACGGATGGCGATATGGA       960
AGGTATTCCGGTAGCGCTGATGGAGGCGATGGCGGTAGGGATTCCCGTGGTATCTACCGT      1020
GCATAGCGGTATTCCGGAACTGGTGGAGGCCGGCAAATCCGGCTGGCTGGTGCCGGAAAA      1080
CGATGCGCAGGCGCTGGCGGCCCGACTCGCTGAGTTCAGCCGGATTGACCACGACACGCT      1140
GGAGTCGGTGATCACGCGCGCCCGTGAAAAAGTGGCGCAAGATTTTAATCAGCAGGCGAT      1200
TAATCGCCAGTTAGCCAGCCTGCTACAAACGATATAAACGAGGTGGTATGCCCGCGACTA      1260
AATTCTCCCGACGTACCCTCCTGACGGCAGGTTCTGCGCTTGCTGTTCTTCCTTTTCTGC      1320
GCGCCTTGCCGGTACAGGCGCGTGAACCTCGCGAGACCGTCGATATTAAGGATTATCCGG      1380
CGGATGACGGTATCGCCTCGTTCAAACAGGCCTTCGCCGACGGACAGACCGTGGTCGTAC      1440
CGCCAGGATGGGTGTGTGAAAATATCAATGCGGCGATAACGATTCCGGCGGGAAAAACGC      1500
TGCGGGTACAGGGCGCGGTGCGTGGGAATGGCCGGGGACGGTTTATTTTGCAGGACGGGT      1560
GTCAGGTGGTGGGGGAGCAGGGCGGCAGTCTGCACAATGTGACGCTGGATGTTCGCGGGT      1620
CGGACTGTGTGATTAAAGGCGTGGCGATGAGCGGCTTTGGCCCCGTCGCGCAAATTTTCA      1680
TCGGTGGTAAGGAACCGCAGGTGATGCGTAATCTCATTATCGATGACATCACCGTTACCC      1740
ACGCCAACTACGCCATTCTCCGCCAGGGATTTCATAACCAAATGGATGGCGCGCGGATTA      1800
CGCATAGCCGCTTTAGCGATTTACAGGGGGACGCCATTGAGTGGAATGTCGCGATTCACG      1860
ACCGCGACATCCTGATTTCCGATCATGTCATCGAACGCATTAATTGTACCAATGGCAAAA      1920
TCAACTGGGGGATCGGCATCGGGCTGGCGGGTAGCACCTATGACAACAGTTATCCTGAAG      1980
```

Figure 10/1

```
ACCAGGCAGTAAAAAACTTTGTGGTGGCCAATATTACCGGATCTGATTGCCGACAGCTTG    2040
TGCACGTAGAAAATGGCAAACATTTCGTCATTCGCAATGTCAAAGCCAAAAACATCACGC    2100
CCGGTTTCAGTAAAAATGCGGGTATTGATAACGCAACGATCGCAATTTATGGCTGTGATA    2160
ATTTCGTCATTGATAATATTGATATGACGAATAGTGCCGGGATGCTCATCGGCTATGGCG    2220
TCGTTAAAGGAAAATACCTGTCAATTCCGCAAAACTTTAAATTAAACGCTATTCGGTTGG    2280
ATAATCGCCAGGTTGCTTATAAATTACGCGGCATTCAAATTTCCTCCGGCAACACCCCCT    2340
CTTTTGTCGCCATCACCAATGTACGGATGACGCGTGCTACGCTGGAACTGCATAATCAAC    2400
CGCAGCACCTCTTTCTGCGCAATATCAACGTGATGCAAACTTCAGCGATTGGCCCGGCGT    2460
TAAAAATGCATTTCGATTTGCGTAAAGATGTACGTGGTCAATTTATGGCCCGCCAGGACA    2520
CGCTGCTTTCCCTCGCTAATGTTCATGCCATCAATGAAAACGGGCAGAGTTCCGTGGATA    2580
TCGACAGGATTAATCACCAAACCGTGAATGTCGAAGCAGTGAATTTTTCGCTGCCGAAGC    2640
GGGGAGGGTAAGTACCGCTATTTTTACGAAAATTCCTGGGAAAAAGTTGTTCATACTTAA    2700
TGTTATGGTGCCGACTAAGACGTAATGTAGAGCGTGCCATCATTATCCCTGGCAGCAGAG    2760
TAATTCATGCTGGCGAAAACAAGCTAAAGAGCTATAATTCAGCAACCATTTTACAGGTGG    2820
AAGAAACAATGATGAATTTGAAAGCAGTTATACCGGTAGCGGGTTTGGGTATGCATATGT    2880
TGCCTGCCACCAAGGCAATCCCAAAAGAGATGCTACCGATCGTCGACAAGCCAATGATTC    2940
AGTACATTGTCGATGAGATTGTGGCTGCAGGGATCAAAGAAATCGTGCTGGTGACTCACG    3000
CGTCTAAAAACGCCGTTGAGAACCACTTCGACACCTCTTATGAACTTGAATCACTTCTTG    3060
AGCAGCGCGTTAAGCGTCAGCTTTTGGCGGAAGTGCAATCTATCTGCCCACCGGGCGTGA    3120
CGATTATGAACGTTCGCCAGGCGCAGCCGTTAGGGCTGGGGCATTCTATTCTGTGCGCGC    3180
GTCCGGTCGTGGGCGATAACCCTTTCATTGTGGTACTCCCGGATATTATTATCGATGATG    3240
CTACCGCCGATCCGCTGCGCTATAACCTTGCGGCGATGGTGGCGCGTTTCAATGAAACGG    3300
GTCGCAGCCAGGTGCTGGCGAAGCGCATGAAAGGTGATTTATCGGAGTATTCCGTTATCC    3360
AGACGAAAGAACCTCTGGATAATGAAGGCAAAGTCAGCCGGATTGTGGAGTTTATCGAAA    3420
AACCGGATCAGCCGCAGACGCTGGATTCCGATTTGATGGCGGTAGGCCGTTATGTGCTTT    3480
CAGCCGACATCTGGGCGGAACTGGAAAGAACCGAACCGGGCGCCTGGGGCCGCATCCAGC    3540
TCACCGATGCCATTGCTGAACTGGCGAAAAAACAGTCGGTTGACGCGATGCTAATGACGG    3600
GTGACAGCTATGACTGCGGTAAAAAAATGGGCTACATGCAGGCATTTGTGAAGTACGGGC    3660
TGCGCAACCTGAAAGAAGGAGCCAAGTTCCGTAAGAGCATAGAGCAGCTTTTGCATGAAT    3720
AAGTATTAACAACCGTGATAAATGGTTGGTGATAAACATAATAACGGCAGTGAACATTCG    3780
AAGCGGCAAGTTGGCTGAAACGAGTGTTGACTGCCGTTTTAGTTTTGTATAAAGGGCTTA    3840
AGTAACAAGGGGTTATCTGGAGCATTTTAATGCTGATTTTATAAGATTAATCCTTGTTTC    3900
CGGATGCAATTAATAAGACAATTAGCGTTTAAGTTTTAGTGAGCTTTGCCCTGCTGGGCG    3960
```

Figure 10/2

UY/423093

```
AGGTTTGCAACAAGTCGATATGTACGCAGTGCACTGGTAGCTGATGAGCCAGGGGCGGTA        4020

GCGTGTGTAACGACTTGAGCAATTAATTTTTATTGGCAAATTAAATACCACATTAAATAC        4080

                   Start of rmlB
                       V  K  I  L  I  T  G  G  A  G  F  I  G  S
GCCTTATGGAATAGAAAAGTGAAGATACTTATTACTGGCGGGGCAGGTTTTATTGGATCA        4140

 A  V  V  R  H  I  I  K  N  T  Q  D  T  V  V  N  I  D  K  L
GCTGTTGTCCGCCATATTATTAAGAATACACAGGACACTGTAGTTAATATTGATAAATTA        4200

   T  Y  A  G  N  L  E  S  L  S  D  I  S  E  S  N  R  Y  N  F
ACCTACGCCGGTAATCTTGAATCCCTTTCTGATATTTCTGAAAGTAATCGCTACAATTTT        4260

   E  H  A  D  I  C  D  S  A  E  I  T  R  I  F  E  Q  Y  Q  P
GAACACGCGGATATTTGTGATTCCGCTGAAATAACGCGTATTTTTGAGCAGTACCAGCCG        4320

   D  A  V  M  H  L  A  A  E  S  H  V  D  R  S  I  T  G  P  A
GACGCGGTGATGCATTTGGCTGCGGAAAGTCATGTGGACCGTTCGATTACCGGGCCAGCA        4380

   A  F  I  E  T  N  I  V  G  T  Y  A  L  L  E  V  A  R  K  Y
GCATTTATTGAAACCAATATCGTCGGCACCTATGCACTTCTTGAAGTTGCGCGTAAATAC        4440

   W  S  A  L  G  E  D  K  K  N  N  F  R  F  H  H  I  S  T  D
TGGTCTGCCCTTGGCGAAGATAAAAAAAATAATTTTCGTTTTCATCATATTTCCACTGAT        4500

   E  V  Y  G  D  L  P  H  P  D  E  V  E  N  S  V  T  L  P  L
GAAGTTTACGGCGATTTACCGCATCCTGATGAAGTTGAAAACAGCGTTACGCTGCCGTTA        4560

   F  T  E  T  T  A  Y  A  P  S  S  P  Y  S  A  S  K  A  S  S
TTTACTGAAACGACGGCATATGCGCCAAGTAGCCCCTATTCTGCGTCAAAAGCATCCAGC        4620

   D  H  L  V  R  A  W  R  R  T  Y  G  L  P  T  I  V  T  N  C
GATCATTTAGTCCGTGCCTGGCGGCGTACCTATGGTCTACCAACGATCGTTACCAATTGT        4680

   S  N  N  Y  G  P  Y  H  F  P  E  K  L  I  P  L  V  I  L  N
TCTAATAACTATGGCCCTTATCACTTCCCTGAAAAACTGATTCCGTTGGTCATTTTGAAC        4740

   A  L  E  G  K  P  L  P  I  Y  G  K  G  D  Q  I  R  D  W  L
GCACTGGAAGGAAAGCCTTTGCCAATTTATGGCAAAGGGGATCAGATTCGCGATTGGCTA        4800

   Y  V  E  D  H  A  R  A  L  H  M  V  V  T  E  G  K  A  G  E
TATGTAGAAGATCATGCTCGCGCGCTTCATATGGTAGTGACTGAAGGCAAGGCAGGGGAG        4860

   T  Y  N  I  G  G  H  N  E  K  K  N  L  D  V  V  F  T  I  C
ACTTATAACATTGGTGGACACAATGAGAAGAAAAATCTCGATGTGGTATTTACCATCTGT        4920

   D  L  L  D  E  I  V  P  K  A  T  S  Y  R  E  Q  I  T  Y  V
GATCTGCTGGATGAGATTGTACCCAAAGCGACTTCTTATCGTGAACAAATCACTTATGTC        4980

   A  D  R  P  G  H  D  R  R  Y  A  I  D  A  G  K  I  S  R  E
GCGGATCGTCCGGGCCATGATCGTCGTTATGCCATTGATGCAGGTAAAATTAGCCGCGAA        5040

   L  G  W  K  P  L  E  T  F  E  S  G  I  R  K  T  V  E  W  Y

TTAGGCTGGAAACCGCTGGAGACCTTTGAAAGCGGTATTCGTAAAACAGTGGAATGGTAC        5100

   L  A  N  T  Q  W  V  N  N  V  K  S  G  A  Y  Q  S  W  I  E
CTTGCAAATACTCAATGGGTAAACAATGTTAAAAGTGGGGCGTATCAGAGTTGGATAGAA        5160

End of rmlB    Start of rmlD
   Q  N  Y  E  G  R  Q  *
                       M  N  I  L  L  F  G  K  T  G  Q  V
CAGAACTATGAAGGACGCCAG TAATGAATATCTTACTTTTTGGTAAGACAGGGCAAGTAG        5220
```

Figure 10/3

09/42~093

```
G   W   E   L   Q   R   S   L   A   P   V   G   N   L   I   A   L   D   V   H
GCTGGGAGTTGCAACGTTCTCTGGCACCGGTAGGGAATCTGATTGCCCTGGATGTCCATT          5280

S   K   E   F   C   G   D   F   S   N   P   K   G   V   A   E   T   V   R   K
CAAAAGAGTTTTGCGGTGATTTTAGTAATCCGAAAGGCGTTGCCGAAACCGTTCGTAAGC          5340

L   R   P   D   V   I   V   N   A   A   A   H   T   A   V   D   K   A   E   S
TTCGTCCCGATGTGATTGTTAACGCAGCAGCCCATACTGCAGTAGATAAAGCAGAGTCTG          5400

E   P   E   L   A   Q   L   L   N   A   T   S   V   E   A   I   A   K   A   A
AACCAGAACTGGCGCAGTTACTTAACGCCACCAGTGTGGAAGCCATCGCTAAAGCAGCCA          5460

N   E   T   G   A   W   V   V   H   Y   S   T   D   Y   V   F   P   G   T   G
ACGAAACTGGCGCATGGGTAGTGCATTATTCAACCGATTATGTATTTCCTGGTACCGGCG          5520

D   I   P   W   Q   E   T   D   A   T   S   P   L   N   V   Y   G   K   T   K
ATATCCCATGGCAGGAAACGGACGCTACGTCGCCGCTGAATGTCTATGGCAAAACCAAAC          5580

L   A   G   E   K   A   L   Q   D   N   C   P   K   H   L   I   F   R   T   S
TGGCGGGAGAAAAGGCCCTGCAGGATAACTGCCCTAAACACCTTATCTTCCGCACCAGTT          5640

W   V   Y   A   G   K   G   N   N   F   A   K   T   M   L   R   L   A   K   E
GGGTTTATGCAGGTAAGGGCAATAATTTCGCAAAGACAATGCTTCGTCTGGCGAAAGAGC          5700

R   Q   T   L   S   V   I   N   D   Q   Y   G   A   P   T   G   A   E   L   L
GTCAGACACTTTCAGTCATTAACGATCAGTACGGTGCGCCAACCGGTGCGGAATTACTGG          5760

A   D   C   T   A   H   A   I   R   V   A   L   N   K   P   E   V   A   G   L
CTGACTGTACGGCGCATGCGATCCGTGTGGCGTTAAATAAACCAGAAGTCGCAGGTCTTT          5820

Y   H   L   V   A   G   G   T   T   T   W   H   D   Y   A   A   L   V   F   D
ACCATCTGGTTGCCGGGGGGAACCACAACCTGGCATGACTACGCGGCCTTAGTCTTTGACG          5880

E   A   R   K   A   G   I   T   L   A   L   T   E   L   N   A   V   P   T   S
AGGCGCGCAAAGCAGGGATAACGCTTGCGCTGACTGAGCTTAATGCTGTGCCGACCAGCG          5940

A   Y   P   T   P   A   S   R   P   G   N   S   R   L   N   T   E   K   F   Q
CCTACCCGACGCCGGCGAGCAGACCAGGCAATTCGCGTCTCAATACTGAAAAGTTTCAGC          6000

R   N   F   D   L   I   L   P   Q   W   E   L   G   V   K   R   M   L   T   E
GTAATTTTGACCTTATTCTGCCTCAATGGGAATTAGGAGTTAAGCGTATGCTGACTGAAA          6060
```

End of rmlD
```
M   F   T   T   T   T   I   *
TGTTTACGACGACAACCATC TAATAAATTTAAATGCCCATCAGGGCATTTTCTATGAATG          6120
```

Start of rmlA
```
            M   K   T   R   K   G   I   I   L   A   G   G   S   G   T   R   L
AGAAATGGAAATGAAAACGCGTAAGGGCATTATTTTAGCGGGGGGCTCCGGCACCCGTCT          6180

    Y   P   V   T   M   A   V   S   K   Q   L   L   P   I   Y   D   K   P   M   I
TTATCCGGTGACCATGGCGGTAAGTAAGCAATTGCTACCAATTTATGATAAACCGATGAT          6240

    Y   Y   P   L   S   T   L   M   L   A   G   I   R   D   I   L   I   I   S   T
TTACTATCCCCTTTCCACGCTTATGCTGGCAGGCATTCGGGATATCCTGATCATCAGTAC          6300

    P   Q   D   T   P   R   F   Q   Q   L   L   G   D   G   S   Q   W   G   L   N
GCCACAGGACACGCCGCGTTTTCAACAACTGCTGGGAGACGGCAGCCAGTGGGGGCTGAA          6360

    L   Q   Y   K   V   Q   P   S   P   D   G   L   A   Q   A   F   I   I   G   E
TCTTCAATATAAAGTACAGCCAAGCCCGGATGGCTTAGCACAGGCGTTTATTATTGGTGA          6420

    E   F   I   G   H   D   D   C   A   L   V   L   G   D   N   I   F   Y   G   H
AGAGTTCATTGGTCATGATGATTGTGCATTAGTGCTGGGTGACAATATCTTCTATGGTCA          6480
```

Figure 10/4

```
         D   L   P   K   L   M   E   A   A   V   N   K   E   S   G   A   T   V   F   A
        TGATTTACCAAAGTTAATGGAAGCTGCCGTTAATAAAGAAAGTGGTGCTACCGTCTTCGC        6540

         Y   H   V   N   D   P   E   R   Y   G   V   V   E   F   D   Q   K   G   T   A
        TTATCATGTAAACGATCCGGAGCGCTACGGTGTGGTTGAGTTTGACCAAAAGGGCACAGC        6600

         V   S   L   E   E   K   P   L   Q   P   K   S   N   Y   A   V   T   G   L   Y
        CGTTAGTCTGGAAGAAAAACCATTACAACCGAAGAGTAATTACGCGGTAACGGGGCTGTA        6660

         F   Y   D   N   S   V   V   E   M   A   K   N   L   K   P   S   A   R   G   E
        TTTTTATGATAATAGCGTGGTGGAGATGGCGAAAAATCTTAAGCCTTCCGCTCGCGGTGA        6720

         L   E   I   T   D   I   N   R   I   Y   M   E   Q   G   R   L   S   V   A   M
        GTTAGAAATCACGGATATTAACCGTATCTATATGGAGCAGGGAAGATTGTCTGTCGCTAT        6780

         M   G   R   G   Y   A   W   L   D   T   G   T   H   Q   S   L   I   E   A   S
        GATGGGCGCGGTTATGCCTGGCTGGATACAGGGACGCATCAGAGTTTGATAGAGGCCAG        6840

         N   F   I   A   T   I   E   E   R   Q   G   L   K   V   S   C   P   E   E   I
        TAATTTTATTGCAACCATCGAAGAACGCCAGGGGCTAAAAGTGTCCTGCCCGGAAGAGAT        6900

         A   F   R   K   N   F   I   N   A   Q   Q   V   I   E   L   A   G   P   L   S
        CGCATTTCGTAAAAATTTTATAAATGCACAACAGGTTATAGAACTGGCCGGGCCATTATC        6960

                                        End of rmlA    Start of rmlC
         K   N   D   Y   G   K   Y   L   L   K   M   V   K   G   L   *   V   M   I   V
        AAAAAATGATTATGGCAAATATTTGCTGAAGATGGTGAAAGGTTTA TAAGTGATGATTGT        7020

         I   K   T   A   I   P   D   V   L   I   L   E   P   K   V   F   G   D   E   R
        GATTAAAACAGCAATACCAGATGTCTTGATCTTAGAGCCTAAAGTTTTTGGCGATGAGAG        7080

         G   F   F   F   E   S   Y   N   Q   Q   T   F   E   E   L   I   G   R   K   V
        GGGATTCTTTTTTGAAAGTTATAACCAGCAGACCTTTGAAGAGTTGATTGGACGTAAAGT        7140

         T   F   V   Q   D   N   H   S   K   S   K   K   N   V   L   R   G   L   H   F
        TACATTTGTTCAAGATAATCATTCAAAATCCAAAAAGAACGTACTCAGAGGGCTACATTT        7200

         Q   R   G   E   N   A   Q   G   K   L   V   R   C   A   V   G   E   V   F   D
        TCAGAGAGGAGAAAATGCACAGGGGAAGTTAGTTCGTTGTGCTGTCGGTGAGGTTTTTGA        7260

         V   A   V   D   I   R   K   E   S   P   T   F   G   Q   W   V   G   V   N   L
        TGTTGCGGTCGATATCCGAAAAGAATCGCCTACTTTTGGTCAATGGGTTGGTGTAAATCT        7320

         S   A   E   N   K   R   Q   L   W   I   P   E   G   F   A   H   G   F   V   T
        GTCTGCTGAGAATAAGCGACAGCTTTGGATTCCAGAAGGTTTTGCTCATGGTTTTGTTAC        7380

         L   S   E   Y   A   E   F   L   Y   K   A   T   N   Y   Y   S   P   S   S   E
        TCTTAGTGAGTATGCAGAGTTTCTGTACAAAGCAACTAATTATTACTCACCTTCATCGGA        7440

         G   S   I   L   W   N   D   E   A   I   G   I   E   W   P   F   S   Q   L   P
        AGGTAGCATTCTATGGAATGATGAGGCAATAGGTATTGAATGGCCTTTTTCTCAGCTGCC        7500

                                                            End of rmlC
         E   L   S   A   K   D   A   A   A   P   L   L   D   Q   A   L   L   T   E   *
        TGAGCTTTCAGCAAAAGATGCTGCAGCACCTTTACTGGATCAAGCCTTGTTAACAGAG TA        7560

         Start of ddhD
                 V   S   H   I   I   K   I   F   P   S   N   I   E   F   S   G   R   E
        AGCATCGTGTCTCATATTATTAAGATTTTTCCATCAAATATTGAATTTTCCGGTAGAGAG        7620

         D   E   S   I   L   D   A   A   L   S   A   G   I   H   L   E   H   S   C   K
        GATGAATCAATCCTCGATGCTGCGCTATCGGCTGGTATCCATCTTGAACATAGCTGCAAA        7680

         A   G   D   C   G   I   C   E   S   D   L   L   A   G   E   V   V   D   S   K
        GCGGGTGATTGTGGTATCTGTGAGTCCGATTTGTTGGCGGGAGAAGTTGTTGACTCCAAA        7740
```

Figure 10/5

```
        G   N   I   F   G   Q   G   D   K   I   L   T   C   C   C   K   P   K   T   A
        GGTAATATTTTTGGACAGGGTGATAAAATACTAACCTGCTGCTGTAAACCTAAAACCGCC          7800

        L   E   L   N   A   H   F   F   P   E   L   A   G   Q   T   K   K   I   V   P
        CTTGAGCTAAATGCGCATTTTTTTCCTGAACTAGCTGGACAGACAAAAAAAATTGTCCCA          7860

        C   K   V   N   S   A   V   L   V   S   G   D   V   M   T   L   K   L   R   T
        TGCAAGGTAAATAGTGCTGTACTGGTTTCAGGCGATGTTATGACTTTGAAGTTACGCACA          7920

        P   P   T   A   K   I   G   F   L   P   G   Q   Y   I   N   L   H   Y   K   G
        CCACCAACAGCAAAAATTGGCTTCCTTCCAGGGCAGTATATCAATTTACATTATAAAGGT          7980

        V   T   R   S   Y   S   I   A   N   S   D   E   S   N   G   I   E   L   H   V
        GTAACTCGCAGTTATTCTATCGCTAATAGTGATGAGTCGAATGGTATTGAGTTGCATGTA          8040

        R   N   V   P   N   G   Q   M   S   S   L   I   F   G   E   L   Q   E   N   T
        AGGAATGTTCCCAATGGTCAGATGAGTTCGCTCATTTTTGGGGAGTTACAAGAAAATACT          8100

        L   M   R   I   E   G   P   C   G   T   F   F   I   R   E   S   D   R   P   I
        CTTATGCGCATTGAAGGGCCTTGCGGAACATTTTTTATTCGTGAAAGTGACAGACCTATA          8160

        I   F   L   A   G   G   T   G   F   A   P   V   K   S   M   V   E   H   L   I
        ATCTTCCTTGCAGGCGGTACTGGATTCGCTCCAGTTAAATCAATGGTTGAGCATCTCATT          8220

        Q   G   K   C   R   R   E   I   Y   I   Y   W   G   M   Q   Y   S   K   D   F
        CAGGGAAAATGTCGTCGTGAGATCTACATTTACTGGGGAATGCAATATAGTAAAGATTTT          8280

        Y   S   A   L   P   Q   Q   W   S   E   Q   H   D   N   V   H   Y   I   P   V
        TACTCTGCATTACCGCAGCAGTGGAGTGAACAGCACGACAACGTTCATTATATCCCTGTT          8340

        V   S   G   D   D   A   E   W   G   G   R   K   G   F   V   H   H   A   V   M
        GTTTCTGGTGATGACGCCGAATGGGGGGGAAGAAAGGGATTTGTCCATCATGCCGTGATG          8400

        D   D   F   D   S   L   E   F   F   D   I   Y   A   C   G   S   P   V   M   I
        GATGATTTTGATTCTCTAGAGTTCTTCGATATATATGCATGTGGTTCACCTGTGATGATC          8460

        D   A   S   K   K   D   F   M   M   K   N   L   S   V   E   H   F   Y   S   D
        GATGCCAGTAAAAAGGACTTTATGATGAAAAATCTCTCTGTAGAACATTTCTATTCTGAT          8520

                                           End of ddhD    Start of ddhA
        A   F   T   A   S   N   N   I   E   D   N   L   *
                                                        M   K   A   V   I   L   A   G
        GCATTTACCGCATCTAATAATATTGAGGATAATTTTATGAAAGCGGTCATCCTGGCTGGTG          8580

        G   L   G   T   R   L   S   E   E   T   I   V   K   P   K   P   M   V   E   I
        GACTTGGTACCAGACTAAGTGAAGAAACAATTGTAAAACCAAAACCGATGGTAGAAATTG          8640

        G   G   K   P   I   L   W   H   I   M   K   M   Y   S   V   H   G   I   K   D
        GTGGCAAGCCTATTCTTTGGCACATTATGAAAATGTATTCTGTGCATGGTATCAAGGATT          8700

        F   I   I   C   C   G   Y   K   G   Y   V   I   K   E   Y   F   A   N   Y   F
        TTATTATCTGCTGTGGTTATAAAGGATATGTGATTAAAGAATATTTTGCGAACTACTTCC          8760

        L   H   M   S   D   V   T   F   H   M   A   E   N   R   M   E   V   H   H   K
        TTCACATGTCAGATGTAACATTCCATATGGCTGAAAACCGTATGGAAGTTCACCATAAAC          8820

        R   V   E   P   W   N   V   T   L   V   D   T   G   D   S   S   M   T   G   G
        GTGTTGAACCATGGAATGTCACATTGGTTGATACGGGTGATTCTTCAATGACTGGTGGTC          8880

        R   L   K   R   V   A   E   Y   V   K   D   D   E   A   F   L   F   T   Y   G
        GTCTGAAACGTGTTGCTGAATACGTAAAAGATGACGAGGCTTTCCTGTTTACTTATGGTG          8940

        D   G   V   A   D   L   D   I   K   A   T   I   D   F   H   K   A   H   G   K
        ATGGCGTTGCCGACCTTGATATCAAAGCGACTATCGATTTCCATAAGGCTCACGGTAAGA          9000
```

Figure 10/6

09/423093

```
K   A   T   L   T   A   T   F   P   P   G   R   F   G   A   L   D   I   R   A
AAGCGACTTTAACAGCTACTTTTCCACCAGGACGCTTTGGCGCATTAGATATCCGAGCTG          9060

G   Q   V   R   S   F   Q   E   K   P   K   G   D   G   A   M   I   N   G   G
GTCAGGTCCGGTCATTCCAGGAAAAACCGAAAGGCGATGGGGCAATGATCAATGGTGGTT          9120

F   F   V   L   N   P   S   V   I   D   L   I   D   N   D   A   T   T   W   E
TCTTTGTGTTGAATCCATCGGTTATCGATCTCATCGATAACGATGCAACAACCTGGGAAC          9180

Q   E   P   L   M   T   L   A   Q   Q   G   E   L   M   A   F   E   H   P   G
AAGAGCCATTAATGACATTGGCACAACAGGGGGAGTTAATGGCTTTTGAACACCCAGGTT          9240

F   W   Q   P   M   D   T   L   R   D   K   V   Y   L   E   G   L   W   E   K
TCTGGCAGCCGATGGATACCCTACGTGATAAAGTTTACCTCGAAGGGCTGTGGGAAAAG          9300

                End of ddhA   Start of ddhB
                                  M   I   D   K   N   F   W   Q   G
G   K   A   P   W   K   T   W   E   *
GTAAAGCTCCGTGGAAAACCTGGGAG TAACTAGATGATTGATAAAAATTTTTGGCAAGGT          9360

K   R   V   F   V   T   G   H   T   G   F   K   G   S   W   L   S   L   W   L
AAACGTGTATTCGTTACCGGCCATACTGGCTTTAAAGGAAGCTGGCTTTCGCTATGGCTG          9420

T   E   M   G   A   I   V   K   G   Y   A   L   D   A   P   T   V   P   S   L
ACTGAAATGGGTGCAATTGTAAAAGGCTATGCACTTGATGCGCCAACTGTTCCAAGTTTA          9480

F   E   I   V   R   L   N   D   L   M   E   S   H   I   G   D   I   R   D   F
TTTGAGATAGTGCGTCTTAATGATCTTATGGAATCTCATATTGGCGACATTCGTGATTTT          9540

E   K   L   R   N   S   I   A   E   F   K   P   E   I   V   F   H   M   A   A
GAAAAGCTGCGCAATTCTATTGCAGAATTTAAGCCAGAAATTGTTTTCCATATGGCAGCC          9600

Q   P   L   V   R   L   S   Y   E   Q   P   I   E   T   Y   S   T   N   V   M
CAGCCTTTAGTGCGCCTATCTTATGAACAGCCAATCGAAACATACTCAACAAATGTTATG          9660

G   T   V   H   L   L   E   T   V   K   Q   V   G   N   I   K   A   V   V   N
GGTACTGTCCATTTGCTTGAAACAGTTAAGCAAGTAGGTAACATAAAGGCAGTCGTAAAT          9720

I   T   S   D   K   C   Y   D   N   R   E   W   V   W   G   Y   R   E   N   E
ATCACCAGTGATAAGTGCTACGACAATCGTGAGTGGGTGTGGGGCTATCGTGAGAACGAA          9780

P   M   G   G   Y   D   P   Y   S   N   S   K   G   C   A   E   L   V   A   S
CCCATGGGAGGGTACGATCCATACTCTAATAGTAAAGGTTGTGCAGAATTAGTCGCGTCT          9840

A   F   R   N   S   F   F   N   P   A   N   Y   E   Q   H   G   V   G   L   A
GCATTCCGGAACTCATTCTTCAATCCTGCAAATTATGAGCAACATGGCGTTGGTTTGGCG          9900

S   V   R   A   G   N   V   I   G   G   G   D   W   A   K   D   R   L   I   P
TCTGTGAGGGCTGGTAATGTCATAGGCGGAGGCGATTGGGCTAAAGACCGTTTAATTCCC          9960

D   I   L   R   S   F   E   N   N   Q   Q   V   I   I   R   N   P   Y   S   I
GATATTCTGCGCTCATTTGAAAATAACCAGCAGGTTATTATTCGAAACCCATATTCTATC         10020

R   P   W   Q   H   V   L   E   P   L   S   G   Y   I   V   V   A   Q   R   L
CGTCCCTGGCAGCATGTACTGGAGCCTCTTTCTGGTTACATTGTGGTGGCGCAACGCTTA         10080

Y   T   E   G   A   K   F   S   E   G   W   N   F   G   P   R   D   E   D   A
TATACAGAAGGTGCTAAGTTTTCTGAAGGATGGAATTTCGGCCCGCGTGATGAAGATGCG         10140

K   T   V   E   F   I   V   D   K   M   V   T   L   W   G   D   D   A   S   W
AAGACGGTCGAATTTATTGTTGACAAGATGGTCACGCTTTGGGGTGATGATGCAAGCTGG         10200

L   L   D   G   E   N   H   P   H   E   A   H   Y   L   K   L   D   C   S   K
TTACTGGATGGTGAGAATCATCCTCATGAGGCACATTACCTGAAACTGGATTGCTCTAAA         10260
```

Figure 10/7

```
       A  N  M  Q  L  G  W  H  P  R  W  G  L  T  E  T  L  G  R  I
       GCAAATATGCAATTAGGATGGCATCCGCGTTGGGGATTGACTGAAACACTTGGTCGCATC      10320

       V  K  W  H  K  A  W  I  R  G  E  D  M  L  I  C  S  K  R  E
       GTAAAATGGCATAAAGCATGGATTCGCGGCGAAGATATGTTGATTTGTTCAAAGCGTGAA      10380

                                       End of ddhB
       I  S  D  Y  M  S  A  T  T     T  R  *
       ATCAGCGACTATATGTCTGCAACTACTCGT TAAGAAAATAAGTTTAAGGAATCAAAGTAA      10440

           Start of ddhC
       M  T  A  N  N  L  R  E  Q  I  S  Q  L  V  A  Q  Y  A  N  E
       TGACAGCAAATAACCTGCGTGAGCAAATCTCTCAGCTTGTCGCTCAGTATGCGAATGAGG      10500

       A  L  S  P  K  P  F  V  A  G  T  S  V  V  P  P  S  G  K  V
       CATTGAGCCCGAAACCTTTTGTTGCAGGTACAAGCGTTGTGCCTCCTTCCGGGAAGGTTA      10560

       I  G  A  K  E  L  Q  L  M  V  E  A  S  L  D  G  W  L  T  T
       TTGGTGCCAAAGAGTTACAATTGATGGTTGAGGCGTCTCTTGATGGATGGCTAACTACTG      10620

       G  R  F  N  D  A  F  E  K  K  L  G  E  F  I  G  V  P  H  V
       GTCGTTTCAATGATGCCTTTGAAAAAAAACTTGGGGAATTTATTGGGGTTCCTCATGTTT      10680

       L  T  T  S  G  S  S  A  N  L  L  A  L  T  A  L  T  S  P
       TAACGACAACATCTGGCTCTTCGGCAAACTTGCTGGCACTGACTGCGCTGACTTCCCCAA      10740

       K  L  G  E  R  A  L  K  P  G  D  E  V  I  T  V  A  A  G  F
       AATTAGGCGAGCGAGCTCTCAAACCTGGTGATGAGGTTATTACTGTCGCTGCTGGCTTCC      10800

       P  T  T  V  N  P  A  I  Q  N  G  L  I  P  V  F  V  D  V  D
       CGACTACAGTTAACCCGGCGATCCAGAATGGTTTAATACCGGTATTCGTGGATGTTGATA      10860

       I  P  T  Y  N  I  D  A  S  L  I  E  A  A  V  T  E  K  S  K
       TCCCGACATATAATATCGATGCCTCTCTCATTGAAGCTGCAGTTACTGAGAAATCAAAAG      10920

       A  I  M  I  A  H  T  L  G  N  A  F  N  L  S  E  V  R  R  I
       CGATAATGATCGCTCATACACTCGGTAATGCATTTAACCTGAGTGAAGTTCGTCGGATTG      10980

       A  D  K  Y  N  L  W  L  I  E  D  C  C  D  A  L  G  T  T  Y
       CCGATAAATATAACTTATGGTTGATTGAAGACTGCTGTGATGCCCTTGGGACGACTTATG      11040

       E  G  Q  M  V  G  T  F  G  D  I  G  T  V  S  F  Y  P  A  H
       AAGGCCAGATGGTAGGTACCTTTGGTGACATCGGAACCGTTAGTTTTTATCCGGCTCACC      11100

       H  I  T  M  G  E  G  G  A  V  F  T  K  S  G  E  L  K  K  I
       ATATCACAATGGGTGAAGGCGGTGCTGTATTCACCAAGTCAGGTGAACTGAAGAAAATTA      11160

       I  E  S  F  R  D  W  G  R  D  C  Y  C  A  P  G  C  D  N  T
       TTGAGTCGTTCCGTGACTGGGGCCGGGATTGTTATTGTGCGCCAGGATGCGATAACACCT      11220

       C  G  K  R  F  G  Q  Q  L  G  S  L  P  Q  G  Y  D  H  K  Y
       GCGGTAAACGTTTTGGTCAGCAATTGGGATCACTTCCTCAAGGCTATGATCACAAATATA      11280

       T  Y  S  H  L  G  Y  N  L  K  I  T  D  M  Q  A  A  C  G  L
       CTTATTCCCACCTCGGATATAATCTCAAAATCACGGACATGCAGGCAGCATGTGGTCTGG      11340

       A  Q  L  E  R  V  E  E  F  V  E  Q  R  K  A  N  F  S  Y  L
       CTCAGTTGGAGCGCGTAGAAGAGTTTGTAGAGCAGCGTAAAGCTAACTTTTCCTATCTGA      11400

       K  Q  G  L  Q  S  C  T  E  F  L  E  L  P  E  A  T  E  K  S
       AACAGGGCTTGCAATCTTGCACTGAATTCCTCGAATTACCAGAAGCAACAGAGAAATCAG      11460

       D  P  S  W  F  G  F  P  I  T  L  K  E  T  S  G  V  N  R  V
       ATCCATCCTGGTTTGGCTTCCCTATCACCCTGAAAGAAACTAGCGGTGTTAACCGTGTCG      11520
```

Figure 10/8

09/423093

```
      E   L   V   K   F   L   D   E   A   K   I   G   T   R   L   L   F   A   G   N
     AACTGGTGAAATTCCTTGATGAAGCAAAAATCGGTACACGTTTACTGTTTGCTGGAAATC        11580

          L   I   R   Q   P   Y   F   A   N   V   K   Y   R   V   V   G   E   L   T   N
     TGATTCGCCAACCGTATTTTGCTAATGTGAAATATCGTGTAGTGGGTGAGTTGACAAATA      11640

      T   D   R   I   M   N   Q   T   F   W   I   G   I   Y   P   G   L   T   T   E
     CCGACCGTATAATGAATCAAACGTTCTGGATTGGTATTTATCCAGGCTTGACTACAGAGC       11700

                                                     End of ddhC
      H   L   D   Y   V   V   S   K   F   E   E   F   F   G   L   N   F   *
     ATTTAGATTATGTAGTTAGCAAGTTTGAAGAGTTCTTTGGTTTGAATTTC TAATTCAATT      11760

         Start of abe
                        M   T   F   L   K   E   Y   V   I   V   S   G   A
     TATTCTATCTGGTGATTGCGATGACCTTTTTGAAAGAATATGTAATTGTCAGTGGGGCTT        11820

      S   G   F   I   G   K   H   L   L   E   A   L   K   K   S   G   I   S   V   V
     CCGGCTTTATTGGTAAGCATTTACTCGAAGCGCTAAAAAAATCGGGGATTTCAGTTGTCG       11880

      A   I   T   R   D   V   I   K   N   N   S   N   A   L   A   N   V   R   W   C
     CAATCACTCGAGATGTAATAAAAAATAATAGTAATGCATTAGCTAATGTTAGATGGTGCA       11940

      S   W   D   N   I   E   L   L   V   E   E   L   S   I   D   S   A   L   I   G
     GTTGGGATAATATCGAATTATTAGTCGAGGAGTTATCAATTGATTCTGCATTAATTGGTA      12000

      I   I   H   L   A   T   E   Y   G   H   K   T   S   S   L   I   N   I   E   D
     TCATTCATTTGGCAACAGAATATGGGCATAAAACATCATCTCTCATAAATATTGAAGATG      12060

      A   N   V   I   K   P   L   K   L   L   D   L   A   I   K   Y   R   A   D   I
     CAAATGTTATAAAACCATTAAAGCTTCTTGATTTGGCAATAAAATATCGGGCGGATATCT      12120

      F   L   N   T   D   S   F   F   A   K   K   D   F   N   Y   Q   H   M   R   P
     TTTTAAATACAGATAGTTTTTTTTGCCAAGAAAGATTTTAATTATCAACATATGCGGCCTT     12180

      Y   I   I   T   K   R   H   F   D   E   I   G   H   Y   Y   A   N   M   H   D
     ATATAATTACTAAAAGACACTTTGATGAAATTGGGCATTATTATGCTAATATGCATGACA       12240

      I   S   F   V   N   M   R   L   E   H   V   Y   G   P   G   D   G   E   N   K
     TTTCATTTGTAAACATGCGATTAGAGCATGTATATGGGCCTGGGGATGGTGAAAATAAAT       12300

      F   I   P   Y   I   I   D   C   L   N   K   K   Q   S   C   V   K   C   T   T
     TTATTCCATACATTATCGACTGCTTAAATAAAAAACAGAGTTGCGTGAAATGTACAACAG      12360

      G   E   Q   I   R   D   F   I   F   V   D   D   V   V   N   A   Y   L   T   I
     GCGAACAGATAAGAGACTTTATTTTTGTAGATGATGTGGTAAATGCTTATTTAACTATAT      12420

      L   E   N   R   K   E   V   P   S   Y   T   E   Y   Q   V   G   T   G   A   G
     TAGAAAATAGAAAAGAAGTACCTTCATATACTGAGTATCAAGTTGGAACTGGTGCTGGGG      12480

      V   S   L   K   D   F   L   V   Y   L   Q   N   T   M   M   P   G   S   S   S
     TAAGTTTGAAAGATTTTCTGGTTTATTTGCAAAATACTATGATGCCAGGTTCATCGAGTA       12540

      I   F   E   F   G   A   I   E   Q   R   D   N   E   I   M   F   S   V   A   N
     TATTTGAATTTGGTGCGATAGAGCAAAGAGATAATGAAATAATGTTCTCTGTAGCAAATA      12600

      N   K   N   L   K   A   M   G   W   K   P   N   F   D   Y   K   K   G   I   E
     ATAAAAATTTAAAAGCAATGGGCTGGAAACCAAATTTCGATTATAAAAAAGGAATTGAAG       12660

                     End of abe
      E   L   L   K   R   L   *
     AACTACTGAAACGGTTATGAGATTTTCATGATCTTTTAATAAATAAATCGTTAACAAATT       12720

                                        Start of wzx
                                           V   K   V   Q   L   L
     AGTCGCGTTATGTTGTAAAAACTAAGTCGTTTAATTGCATAGTGAAAGTTCAATTGTTAA       12780
```

Figure 10/9

```
K  I  P  S  H  L  I  V  A  G  S  S  W  L  S  K  I  I  I  A
AAATTCCGAGTCATTTAATTGTTGCAGGTTCATCATGGTTATCCAAAATAATAATTGCCG    12840

G  V  Q  L  A  S  I  S  Y  L  I  S  M  L  G  E  E  K  Y  A
GGGTGCAGTTAGCAAGTATTTCATATCTTATTTCTATGCTAGGTGAAGAGAAATATGCAA    12900

I  F  S  L  L  T  G  L  L  V  W  C  S  A  V  D  F  G  I  G
TCTTTAGTTTGTTAACTGGTTTATTAGTATGGTGTAGCGCTGTTGATTTTGGCATAGGTA    12960

T  G  L  Q  N  Y  I  S  E  C  R  A  K  N  K  S  Y  D  A  Y
CAGGACTGCAAAATTATATATCAGAATGCAGAGCCAAAAACAAAAGTTATGATGCATATA    13020

I  K  S  A  L  H  L  S  F  I  A  I  I  F  F  I  A  L  F  Y
TTAAATCAGCATTACATCTAAGCTTTATAGCTATTATTTTTTTTATTGCTTTATTTTATA    13080

I  F  S  G  V  I  S  A  K  Y  L  S  S  F  H  E  V  L  Q  D
TTTTTTCTGGGGTAATTTCCGCTAAATATCTTTCTTCTTTTCATGAGGTATTACAGGACA    13140

K  T  R  M  L  F  F  T  S  C  L  V  F  S  S  I  G  I  G  A
AAACCAGAATGCTCTTTTTTACCTCATGTCTGGTTTTCAGTTCTATTGGAATCGGAGCTA    13200

I  A  Y  K  I  L  F  A  E  L  V  G  W  K  A  N  L  L  N  A
TTGCTTATAAAATACTTTTTGCCGAATTGGTCGGGTGGAAAGCTAATCTATTAAACGCAT    13260

L  S  Y  M  I  G  M  L  G  L  L  Y  I  Y  Y  R  G  I  S  V
TATCTTATATGATAGGTATGCTCGGCTTGCTATATATATACTATAGGGGGATCTCAGTTG    13320

D  I  K  L  S  L  I  V  L  Y  L  P  V  G  M  I  S  L  C  Y
ACATAAAATTATCACTAATAGTCCTGTATCTTCCAGTGGGTATGATTTCATTGTGCTATA    13380

I  V  Y  R  Y  I  K  L  Y  H  V  K  T  T  K  S  H  Y  I  A
TTGTATATAGATACATAAAGCTTTATCATGTTAAAACAACAAAATCTCATTATATAGCAA    13440

I  L  R  R  S  S  G  F  F  L  F  T  L  L  S  I  V  V  L  Q
TTTTACGTAGATCTTCAGGGTTTTTTCTTTTTACTTTATTATCGATAGTGGTGCTTCAAA    13500

T  D  Y  M  V  I  S  Q  R  L  T  P  A  D  I  V  Q  Y  T  V
CAGATTATATGGTCATTTCTCAAAGGCTAACTCCTGCTGATATTGTTCAATATACAGTAA    13560

T  M  K  I  F  G  L  V  F  F  I  Y  T  A  I  L  Q  A  L  W
CGATGAAAATTTTTGGTTTAGTCTTTTTTATTTATACTGCTATTTTGCAAGCATTATGGC    13620

P  I  C  A  E  L  R  V  K  Q  Q  W  K  K  L  N  K  M  I  G
CTATATGTGCTGAATTGAGAGTCAAACAGCAATGGAAAAAACTTAACAAAATGATAGGTG    13680

V  N  I  L  L  G  S  L  Y  V  V  G  C  T  I  F  I  Y  L  F
TCAATATTTTGCTTGGCTCACTATATGTTGTTGGATGTACAATATTTATTTATTTATTTA    13740

K  E  Q  I  F  S  V  I  A  K  D  I  N  Y  Q  V  S  I  L  S
AAGAACAGATATTTTCAGTAATAGCCAAAGATATTAATTATCAAGTTTCTATTTTATCTT    13800

F  M  L  I  G  I  Y  F  C  I  R  V  W  C  D  T  Y  A  M  L
TTATGTTAATTGGCATATATTTCTGTATTCGCGTTTGGTGTGACACTTATGCAATGTTAT    13860

L  Q  S  M  N  Y  L  K  I  L  W  I  L  V  P  L  Q  A  I  I
TGCAAAGTATGAATTATTTAAAAATACTTTGGATATTAGTACCACTACAAGCAATAATTG    13920

G  G  I  A  Q  W  Y  F  S  S  T  L  G  I  S  G  V  L  L  G
GTGGAATAGCACAATGGTATTTTTCTAGTACGCTTGGAATCAGTGGAGTGCTGCTTGGCT    13980

L  I  I  S  F  A  L  T  V  F  W  G  L  P  L  T  Y  L  I  K
TGATTATATCTTTTGCTTTAACTGTTTTTTGGGGGCTTCCACTAACTTACTTAATTAAGG    14040
```

Figure 10/10

End of *wzx* **Start of** *wbaV*

```
A  N  K  G  *      M  L  I  S  F  C  I  P  T  Y  N  R  K  Q
CAAATAAGGGA TAATCAT ATG CTTATATCATTTTGTATTCCAACTTATAATAGAAAACAA   14100

    Y  L  E  E  L  L  N  S  I  N  N  Q  E  K  F  N  L  D  I  E
TATCTTGAAGAGTTGTTGAATAGTATAAATAATCAGGAAAAATTTAATTTAGATATTGAG       14160

    I  C  I  S  D  N  A  S  T  D  G  T  E  E  M  I  D  V  W  R
ATATGTATATCAGATAATGCCTCTACTGATGGTACAGAGGAAATGATTGATGTTTGGAGG       14220

    N  N  Y  N  F  P  I  I  Y  R  R  N  S  V  N  L  G  P  D  R
AACAATTATAATTTCCCAATAATATATCGGCGTAATAGCGTTAACCTTGGGCCAGATAGG       14280

    N  F  L  A  S  V  S  L  A  N  G  D  Y  C  W  I  F  G  S  D
AATTTTCTTGCTTCAGTATCCCTTGCGAATGGGGATTATTGTTGGATATTTGGCAGTGAT       14340

    D  A  L  A  K  D  S  L  A  I  L  Q  T  Y  L  D  S  Q  A  D
GATGCTCTTGCGAAAGACTCGTTAGCGATATTACAAACTTATCTCGATTCTCAAGCAGAT       14400

    I  Y  L  C  D  R  K  E  T  G  C  D  L  V  E  I  R  N  P  H
ATATATTTATGTGACAGAAAAGAGACCGGGTGTGATTTAGTTGAGATTAGAAACCCTCAT       14460

    R  S  W  L  R  T  D  D  E  L  Y  V  F  N  N  N  L  D  R  E
CGTTCTTGGCTCAGAACAGATGATGAACTTTATGTGTTTAATAATAATTTAGATAGGGAA       14520

    I  Y  L  S  R  C  L  S  I  G  G  V  F  S  Y  L  S  S  L  I
ATCTATCTCAGTAGATGCTTATCTATTGGTGGTGTATTTAGCTATCTAAGTTCTTTAATA       14580

    V  K  K  E  R  W  D  A  I  D  F  D  A  S  Y  I  G  T  S  Y
GTAAAAAAAGAACGATGGGATGCCATTGATTTTGATGCGTCCTATATTGGCACTTCCTAT       14640

    P  H  V  F  I  M  M  S  V  F  N  T  P  G  C  L  L  H  Y  I
CCTCATGTATTTATCATGATGAGCGTATTTAATACGCCAGGGTGCCTTTTGCATTATATA       14700

    S  K  P  L  V  I  C  R  G  D  N  D  S  F  E  K  K  G  K  A
TCAAAACCACTCGTAATATGCCGAGGAGATAATGATAGTTTCGAGAAGAAAGGAAAGGCC       14760

    R  R  I  L  I  D  F  I  A  Y  L  K  L  A  N  D  F  Y  S  K
AGACGAATTTTAATTGATTTTATTGCATATTTAAAATTAGCTAATGATTTTTACAGTAAA       14820

    N  I  S  L  K  R  A  F  E  N  V  L  L  K  E  R  P  W  L  Y
AATATATCTTTAAAACGAGCATTTGAAAATGTTTTGCTAAAAGAGAGACCATGGTTATAT       14880

    T  T  L  A  M  A  C  Y  G  N  S  D  E  K  R  D  L  S  E  F
ACAACTTTGGCTATGGCATGTTATGGCAATAGTGATGAAAAAGAGATTTATCTGAATTT       14940

    Y  A  K  L  G  C  N  K  N  M  I  N  T  V  L  R  F  G  K  L
TATGCAAAGCTAGGTTGTAATAAAAATATGATCAACACTGTACTTCGATTTGGGAAACTA       15000
```

                                                       End of *wbaV*
```
    A  Y  A  V  K  N  I  T  V  L  K  N  F  T  K  R  I  I  K  *
GCATATGCAGTGAAAAATATTACCGTGCTTAAGAATTTTACTAAACGGATAATTAAG TAG     15060

TAGTAAGTTATTATATTGAGATTAAATGTAGATTTAACCTTTCTGGATTCAGCTAGATTT     15120

ACGTTACTGACTTTTCTTTTTAATGAAAATCATATTTGATATATATAAATAAATTTGGAT     15180

AGCTTAACTACTTAGATGTTTTTTTTCTGGGAATGTTAGTATAATAATATATTTCTTTATG     15240

ATTGTTTTTGTAGTGTTTTACTGCCGGTATTACATTAACTCTATTATTAAGAATTACACC    15300

TAGTGTAAGCTTCGTAATATTATTTATCCTTATGATTATTGCTTTAAAGATGCGTATGGA     15360
```

                  Start of *wbaU*
```
                  M  I  V  N  L  S  R  L  G  K  S  G  T  G
AAAACGGAGAGCTATTCA ATG ATCGTAAACCTATCACGTTTAGGTAAAAGTGGTACGGGA    15420
```

Figure 10/11

```
     M   W   Q   Y   S   I   K   F   L   T   A   L   R   E   I   A   D   V   D   A
     ATGTGGCAATACTCGATTAAATTTTTAACGGCACTGCGAGAAATAGCTGATGTTGACGCA         15480

     I   I   C   S   K   V   H   A   D   Y   F   E   K   L   G   Y   A   V   V   T
     ATAATCTGTAGCAAGGTACACGCTGATTATTTTGAAAAGCTCGGTTATGCAGTAGTTACT         15540

     V   P   N   I   V   S   N   T   S   K   T   S   R   L   R   P   L   V   W   Y
     GTTCCGAATATTGTTAGCAACACATCAAAAACATCGCGACTTAGACCATTAGTATGGTAT         15600

     V   Y   S   Y   W   L   A   L   R   V   L   I   K   F   G   N   K   K   L   V
     GTATATAGTTACTGGCTTGCGCTGAGGGTTTTAATTAAGTTTGGTAATAAAAAATTGGTG         15660

     C   T   T   H   H   T   I   P   L   L   R   N   Q   T   I   T   V   H   D   I
     TGTACTACACATCACACTATCCCCTTACTGAGAAACCAAACGATAACCGTACATGATATA         15720

     R   P   F   Y   Y   P   D   S   F   I   Q   K   V   Y   F   R   F   L   L   K
     AGACCTTTTTATTATCCAGATAGTTTTATTCAGAAAGTGTATTTTCGCTTTTTATTAAAA         15780

     M   S   V   K   R   C   K   H   V   L   T   V   S   Y   T   V   K   D   S   I
     ATGTCCGTTAAGCGATGTAAGCATGTTTTAACGGTATCTTATACCGTTAAAGATAGCATT         15840

     A   K   T   Y   N   V   D   S   E   K   I   S   V   I   Y   N   S   V   N   K
     GCTAAAACTTATAATGTAGATAGTGAGAAAATATCAGTAATTTATAATAGTGTTAATAAA        '15900

     S   D   F   I   Q   K   K   E   K   E   N   Y   F   L   A   V   G   A   S   W
     TCTGATTTTATACAAAAAAAAGAAAAAGAGAATTACTTTTTAGCTGTTGGTGCAAGTTGG         15960

     P   H   K   N   I   H   S   F   I   K   N   K   K   V   W   S   D   S   Y   N
     CCACATAAAAATATTCATTCATTCATAAAAAATAAAAAAGTTTGGTCTGACTCTTATAAT         16020

     L   I   I   V   C   G   R   T   D   Y   A   M   S   L   Q   Q   M   V   V   D
     TTAATTATTGTATGTGGTCGTACTGACTATGCAATGTCTCTCCAACAAATGGTCGTTGAT         16080

     L   E   L   K   D   K   V   T   F   L   H   E   V   S   F   N   E   L   K   I
     CTGGAACTAAAAGATAAAGTGACTTTTTTACATGAAGTCTCATTTAATGAATTAAAGATT         16140

     L   Y   S   K   A   Y   A   L   V   Y   P   S   I   D   E   G   F   G   I   P
     TTATATTCTAAAGCCTACGCGCTTGTTTATCCATCTATTGATGAGGGTTTTGGTATACCT         16200

     P   I   E   A   M   A   S   N   T   P   V   I   V   S   D   I   P   V   F   H
     CCTATTGAAGCGATGGCATCAAATACTCCAGTTATAGTGTCCGATATACCAGTATTTCAT         16260

     E   V   L   T   N   G   A   L   Y   V   N   P   D   D   E   K   S   W   Q   S
     GAAGTGTTAACCAATGGTGCATTATATGTGAATCCGGATGATGAAAAAAGCTGGCAGAGT         16320

     A   I   K   N   I   E   Q   L   P   D   A   I   S   R   F   N   N   Y   V   A
     GCAATTAAAAATATAGAGCAGTTGCCTGATGCAATTTCCCGATTTAACAACTATGTCGCA         16380

     R   Y   D   F   D   N   M   K   Q   M   V   G   N   W   L   A   E   S   K   *
     CGGTATGACTTTGATAATATGAAGCAGATGGTTGGCAATTGGTTGGCGGAATCAAAA TAA        16440
```

End of wbaU

Start of wbaN

```
     M   K   I   T   L   I   I   P   T   Y   N   A   G   S   L   W   P   N   V   L
     ATGAAAATAACATTAATTATTCCCACATATAATGCAGGGTCGCTTTGGCCTAATGTTCTG         16500

     D   A   I   K   Q   Q   T   I   Y   P   D   K   L   I   V   I   D   S   G   S
     GATGCGATTAAGCAGCAAACTATATATCCGGATAAATTGATTGTTATAGACTCAGGTTCT         16560

     K   D   E   T   V   P   L   A   S   D   L   K   N   I   S   I   F   N   I   D
     AAAGATGAAACGGTTCCGTTAGCCTCAGACCTGAAAAATATATCAATATTTAATATTGAC         16620

     S   K   D   F   N   H   G   G   T   R   N   L   A   V   A   K   T   L   D   A
     TCTAAAGATTTTAATCATGGAGGAACCAGAAATTTAGCAGTTGCAAAAACTCTGGACGCT         16680
```

Figure 10/12

```
      D   V   I   I   F   L   T   Q   D   A   I   L   A   D   S   D   A   I   K   N
     GATGTTATAATTTTTCTAACGCAAGATGCAATTCTCGCGGATTCGGATGCAATTAAAAAT         16740

      L   V   Y   Y   F   S   D   P   L   I   A   A   V   C   G   R   Q   L   P   H
     TTGGTTTATTATTTTTCAGATCCATTGATAGCAGCGGTTTGTGGTAGACAACTTCCTCAT         16800

      K   D   A   N   P   L   A   V   H   A   R   N   F   N   Y   S   S   K   S   I
     AAAGATGCTAATCCCCTTGCAGTGCATGCCAGAAATTTTAATTATAGTTCAAAATCTATT        16860

      V   K   S   K   A   D   I   E   K   L   G   I   K   T   V   F   M   S   N   S
     GTTAAAAGTAAGGCAGATATAGAAAAATTGGGTATTAAAACTGTATTTATGTCCAATTCT         16920

      F   A   A   Y   R   R   S   V   F   E   E   L   S   G   F   P   E   H   T   I
     TTTGCTGCCTATCGCCGTTCCGTTTTTGAAGAGTTAAGTGGGTTTCCTGAACATACAATT         16980

      L   A   E   D   M   F   M   A   A   K   M   I   Q   A   G   Y   K   V   A   Y
     CTTGCCGAGGATATGTTTATGGCGGCTAAGATGATTCAGGCGGGTTATAAGGTCGCCTAC         17040

      C   A   E   A   V   V   R   H   S   H   N   Y   T   P   R   E   E   F   Q   R
     TGCGCTGAAGCGGTGGTAAGACACTCCCATAATTATACCCCGCGAGAAGAGTTTCAACGA         17100

      Y   F   D   T   G   V   F   H   A   C   S   P   W   I   Q   R   D   F   G   G
     TATTTTGATACTGGTGTATTTCATGCTTGTTCTCCGTGGATTCAGCGTGACTTTGGCGGA         17160

      A   G   G   E   G   F   R   F   V   K   S   E   I   Q   F   L   L   K   N   A
     GCCGGTGGTGAGGGTTTCCGCTTCGTAAAATCAGAGATTCAATTCCTGCTTAAAAATGCA         17220

      P   F   W   I   P   R   A   L   L   T   T   F   A   K   F   L   G   Y   K   L
     CCGTTCTGGATTCCAAGAGCTTTATTAACAACCTTTGCTAAATTCTTGGGTTACAAATTA         17280

      G   K   H   W   Q   S   L   P   L   S   T   C   R   Y   F   S   M   Y   K   S
     GGCAAGCATTGGCAATCTTTACCGTTGTCTACATGTCGCTATTTTAGCATGTACAAGAGT         17340

      Y   W   N   N   I   Q   Y   S   S   S   K   E   I   K   *   M   S   F   L   P
     TATTGGAATAATATCCAATATTCTTCGTCAAAAGAGATAAAA TAAATGTCTTTTCTTCCC          17400
                                            End of wbaN  Start of manC
      V   I   M   A   G   G   T   G   S   R   L   W   P   L   S   R   E   Y   H   P
     GTAATTATGGCTGGCGGCACAGGTAGCCGTTTATGGCCGCTTTCACGCGAATATCATCCG          17460

      K   Q   F   L   S   V   E   G   K   L   S   M   L   Q   N   T   I   K   R   L
     AAGCAGTTTCTAAGCGTTGAAGGTAAACTATCAATGCTGCAAAATACTATAAAGCGATTA          17520

      A   S   L   S   T   E   E   P   V   V   I   C   N   D   R   H   R   F   L   V
     GCTTCACTTTCTACAGAAGAACCCGTTGTCATTTGCAATGACAGACACCGTTTCTTAGTC          17580

      A   E   Q   L   R   E   I   D   K   L   A   N   N   I   I   L   E   P   V   G
     GCTGAACAACTCCGTGAAATTGACAAGTTAGCAAATAATATTATTCTCGAACCGGTAGGC          17640

      R   N   T   A   P   A   I   A   L   A   A   F   C   A   L   Q   N   A   D   N
     CGTAATACTGCACCAGCGATCGCTCTTGCCGCGTTTTGTGCGCTCCAGAATGCTGATAAT          17700

      A   D   P   L   L   L   V   L   A   A   D   H   V   I   Q   D   E   I   A   F
     GCTGATCCTCTTTTGTTGGTTCTTGCTGCAGATCATGTGATTCAGGATGAAATAGCTTTT          17760

      T   K   A   V   R   H   A   E   E   Y   A   A   N   G   K   L   V   T   F   G
     ACGAAAGCTGTCAGACATGCTGAAGAATACGCTGCAAATGGTAAGCTTGTAACTTTTGGT          17820

      I   V   P   T   H   A   E   T   G   Y   G   Y   I   R   R   G   E   L   I   G
     ATTGTTCCAACGCATGCTGAAACGGGTTATGGATATATTCGTCGTGGTGAGTTGATAGGA          17880

      N   D   A   Y   A   V   A   E   F   V   E   K   P   D   I   D   T   A   G   D
     AATGACGCTTATGCAGTGGCTGAATTTGTGGAGAAACCGGATATCGATACCGCCGGTGAC          17940

      Y   F   K   S   G   K   Y   Y   W   N   S   G   M   F   L   F   R   A   S   S
     TATTTCAAATCAGGGAAATATTACTGGAATAGCGGTATGTTTTTATTTCGTGCAAGCTCT          18000
```

Figure 10/13

```
 Y  L  N  E  L  K  Y  L  S  P  E  I  Y  K  A  C  E  K  A  V
TATTTAAACGAATTAAAGTATTTATCACCTGAAATTTATAAAGCTTGTGAAAAGGCGGTA      18060

 G  H  I  N  P  D  L  D  F  I  R  I  D  K  E  E  F  M  S  C
GGACATATAAATCCCGATCTTGATTTTATTCGTATTGATAAAGAAGAGTTTATGTCATGC      18120

 P  S  D  S  I  D  Y  A  V  M  E  H  T  Q  H  A  V  V  I  P
CCGAGTGATTCTATCGATTATGCAGTTATGGAGCACACACAGCATGCGGTGGTGATACCA      18180

 M  S  A  G  W  S  D  V  G  S  W  S  S  L  W  D  I  S  N  K
ATGAGCGCTGGCTGGTCGGATGTGGGTTCCTGGTCCTCACTTTGGGATATATCGAATAAA      18240

 D  H  Q  R  N  V  L  K  G  D  I  F  A  H  A  C  N  D  N  Y
GATCATCAGAGAAATGTTTTAAAAGGAGATATTTTCGCACATGCTTGTAATGATAATTAC      18300

 I  Y  S  E  D  M  F  I  S  A  I  G  V  S  N  L  V  I  V  Q
ATTTATTCCGAAGATATGTTTATAAGTGCGATTGGTGTAAGCAATCTTGTCATTGTTCAA      18360

 T  T  D  A  L  L  V  A  N  K  D  T  V  Q  D  V  K  K  I  V
ACAACAGACGCTTTACTGGTGGCTAATAAAGATACAGTACAAGATGTTAAAAAAATTGTC      18420

 D  Y  L  K  R  N  D  R  N  E  Y  K  Q  H  Q  E  V  F  R  P
GATTATTTAAAACGGAATGATAGGAACGAATATAAACAACATCAAGAAGTTTTCCGCCCC      18480

 W  G  K  Y  N  V  I  D  S  G  K  N  Y  L  V  R  C  I  T  V
TGGGGAAAATATAATGTGATTGATAGCGGCAAAAATTACCTCGTTCGATGTATCACTGTT      18540

 K  P  G  E  K  F  V  A  Q  M  H  H  H  R  A  E  H  W  I  V
AAGCCGGGTGAGAAATTTGTGGCGCAGATGCATCACCACCGGGCTGAGCATTGGATAGTA      18600

 L  S  G  T  A  R  V  T  K  G  E  Q  T  Y  M  V  S  E  N  E
TTATCCGGGACTGCTCGTGTTACAAAGGGAGAGCAGACTTATATGGTTTCTGAAAATGAA      18660

 S  T  F  I  P  P  N  T  I  H  A  L  E  N  P  G  M  T  P  L
TCAACATTTATTCCTCCGAATACTATTCACGCGCTGGAAAATCCTGGAATGACCCCCCTG      18720

 K  L  I  E  I  Q  S  G  T  Y  L  G  E  D  D  I  I  R  L  E
AAGTTAATTGAGATTCAATCAGGTACCTATCTTGGTGAGGATGATATTATTCGTTTAGAA      18780

                        Start of manB    End of manC
                         M  N  V  V  N  N  S  R  D  V
 Q  R  S  G  F  S  K  E  W  T  N  E  R  S  *
CAACGTTCTGGATTTTCGAAGGAGTGGACTAATGAACGTAGTTAATAATAGCCGTGATGT      18840

 I  Y  S  S  G  I  V  F  G  T  S  G  A  R  G  L  V  K  D  F
TATTTATTCATCAGGTATTGTGTTTGGAACGAGTGGGGCTCGCGGTCTTGTAAAAGATTT      18900

 T  P  Q  V  C  A  A  F  T  V  S  F  V  A  V  M  Q  E  H  F
TACACCTCAGGTATGTGCTGCTTTTACGGTTTCATTTGTTGCCGTTATGCAGGAACATTT      18960

 S  F  D  T  V  A  L  A  I  D  N  R  P  S  S  Y  G  M  A  Q
TTCCTTTGATACCGTAGCATTGGCAATAGATAATCGTCCAAGTAGTTATGGGATGGCTCA      19020

 A  C  A  A  A  L  A  D  K  G  V  N  C  I  F  Y  G  V  V  P
GGCGTGTGCTGCTGCATTGGCGGATAAAGGCGTTAACTGTATTTTTATGGAGTGGTACC      19080

 T  P  A  L  A  F  Q  S  M  S  D  N  M  P  A  I  M  V  T  G
AACCCCAGCTTTGGCCTTTCAGTCTATGTCTGACAATATGCCTGCGATAATGGTTACGGG      19140

 S  H  I  P  F  E  R  N  G  L  K  F  Y  R  P  D  G  E  I  T
AAGTCATATTCCATTCGAGCGGAACGGCCTCAAGTTTTATCGTCCTGATGGTGAAATCAC      19200

 K  H  D  E  A  A  I  L  S  V  E  D  T  C  S  H  L  E  L  K
GAAACATGATGAGGCTGCGATCCTTAGTGTTGAAGATACGTGCAGCCATTTAGAGCTTAA      19260
```

Figure 10/14

```
     E   L   I   V   S   E   M   A   A   V   N   Y   I   S   R   Y   T   S   L   F
AGAACTCATAGTTTCAGAAATGGCTGCTGTTAATTATATATCTCGTTATACATCTTTATT        19320

         S   T   P   F   L   K   N   K   R   I   G   I   Y   E   H   S   S   A   G   R
TTCTACTCCATTCCTGAAAAATAAGCGTATTGGTATTTACGAACATTCAAGCGCTGGGCG        19380

     D   L   Y   K   P   L   F   I   A   L   G   A   E   V   V   S   L   G   R   S
TGATCTTTATAAGCCTTTATTTATTGCATTGGGGGCTGAAGTCGTTAGCTTGGGTAGAAG        19440

     D   N   F   V   P   I   D   T   E   A   V   S   K   E   D   R   E   K   A   R
CGATAATTTTGTACCTATAGATACAGAGGCTGTAAGCAAAGAGGATCGGGAAAAAGCTCG        19500

     S   W   A   K   E   F   D   L   D   A   I   F   S   T   D   G   D   G   D   R
CTCATGGGCTAAAGAGTTCGATTTAGATGCCATATTCTCGACAGATGGGGATGGTGATCG        19560

     P   L   I   A   D   E   A   G   E   W   L   R   G   D   I   L   G   L   L   C
CCCTCTTATTGCTGATGAGGCCGGTGAGTGGCTAAGAGGCGATATACTAGGTCTATTATG        19620

     S   L   A   L   D   A   E   A   V   A   I   P   V   S   C   N   S   I   I   S
TTCACTTGCATTGGATGCAGAAGCCGTCGCTATTCCTGTTAGTTGTAACAGCATAATTTC        19680

     S   G   R   F   F   K   H   V   K   L   T   K   I   G   S   P   Y   V   I   E
TTCTGGCCGCTTTTTTAAACATGTTAAGCTTACAAAAATTGGCTCGCCTTATGTTATCGA        19740

     A   F   N   E   L   S   R   S   Y   S   R   I   V   G   F   E   A   N   G   G
AGCTTTTAATGAATTATCGCGGAGTTATAGTCGTATTGTCGGTTTTGAAGCCAATGGCGG        19800

     F   L   L   G   S   D   I   C   I   N   E   Q   N   L   H   A   L   P   T   R
TTTTTTATTAGGAAGCGACATCTGTATTAACGAGCAGAATCTTCATGCCTTACCAACTCG        19860

     D   A   V   L   P   A   I   M   L   L   Y   K   S   R   N   T   S   I   S   A
TGATGCTGTATTACCAGCAATAATGCTGCTTTACAAAAGTAGGAATACCAGCATTAGCGC        19920

     L   V   N   E   L   P   T   R   Y   T   H   S   D   R   L   Q   G   I   T   T
TTTAGTCAATGAACTCCCAACTCGTTACACCCATTCTGACAGATTACAGGGGATTACAAC        19980

     D   K   S   Q   S   L   I   S   M   G   R   E   N   L   S   N   L   L   S   Y
TGATAAAAGTCAATCCTTAATTAGTATGGGCAGAGAAAATCTGAGCAACCTCTTAAGCTA        20040

     I   G   L   E   N   E   G   A   I   S   T   D   M   T   D   G   M   R   I   T
TATTGGTTTGGAGAATGAAGGTGCAATTTCTACAGATATGACAGATGGTATGCGAATTAC        20100

     L   R   D   G   C   I   V   H   L   R   A   S   G   N   A   P   E   L   R   C
TTTACGTGATGGATGTATTGTGCATTTGCGCGCTTCTGGTAATGCACCTGAGTTACGCTG        20160

     Y   A   E   A   N   L   L   N   R   A   Q   D   L   V   N   T   T   L   A   N
CTATGCAGAAGCTAATTTATTAAATAGGGCTCAGGATCTTGTAAATACAACGCTTGCTAA        20220
```

End of manB
```
     I   K   K   R   C   L   L   *
TATTAAAAAACGATGCTTGCTGTAAAAAAATTGAATGTTATTTACTTAATATGCCTATTT        20280
```

Start of wbaP
```
                                          M   D   N   I   D   N   K   Y
TATTTACATTATGCACGGTCAGAGGGTGAGGATTAAATGGATAATATTGATAATAAGTAT        20340

     N   P   Q   L   C   K   I   F   L   A   I   S   D   L   I   F   F   N   L   A
AATCCACAGCTATGTAAAATTTTTTTGGCTATATCGGATTTGATTTTTTTTAATTTAGCC        20400

     L   W   F   S   L   G   C   V   Y   F   I   F   D   Q   V   Q   R   F   I   P
TTATGGTTTTCATTAGGATGTGTCTATTTTATTTTTGATCAAGTACAGCGATTTATTCCT        20460

     Q   D   Q   L   D   T   R   V   I   T   H   F   I   L   S   V   V   C   V   G
CAAGACCAATTAGATACAAGAGTTATTACGCATTTTATTTTGTCAGTAGTATGTGTCGGT        20520
```

Figure 10/15

```
     W   F   W   I   R   L   R   H   Y   T   I   R   K   P   F   W   Y   E   L   K
     TGGTTTTGGATTCGTTTGCGACATTATACTATCCGCAAGCCATTTTGGTATGAGTTAAAA              20580

     E   I   F   R   T   I   V   I   F   A   I   F   D   L   A   L   I   A   F   T
     GAAATTTTTCGTACGATCGTTATTTTTGCTATATTTGATTTGGCTCTGATAGCGTTACA               20640

     K   W   Q   F   S   R   Y   V   W   V   F   C   W   T   F   A   L   I   L   V
     AAATGGCAGTTTTCACGCTATGTCTGGGTGTTTTGTTGGACTTTTGCCCTAATCCTGGTG              20700

     P   F   F   R   A   L   T   K   H   L   L   N   K   L   G   I   W   K   K   K
     CCTTTTTTTCGCGCACTTACAAAGCATTTATTGAACAAGCTAGGTATCTGGAAGAAAAA               20760

     T   I   I   L   G   S   G   Q   N   A   R   G   A   Y   S   A   L   Q   S   E
     ACTATCATCCTGGGGAGCGGACAGAATGCTCGTGGTGCATATTCTGCGCTGCAAAGTGAG              20820

     E   M   M   G   F   D   V   I   A   F   F   D   T   D   A   S   D   A   E   I
     GAGATGATGGGGTTTGATGTTATCGCTTTTTTTGATACGGATGCGTCAGATGCTGAAATA              20880

     N   M   L   P   V   I   K   D   T   E   I   I   W   D   L   N   R   T   G   D
     AATATGTTGCCGGTGATAAAGGATACTGAGATTATTTGGGATTTAAATCGTACAGGTGAT              20940

     V   H   Y   I   L   A   Y   E   Y   T   E   L   E   K   T   H   F   W   L   R
     GTCCATTATATCCTTGCTTATGAATACACCGAGTTGGAGAAAACACATTTTTGGCTACGT              21000

     E   L   S   K   H   H   C   R   S   V   T   V   V   P   S   F   R   G   L   P
     GAACTTTCAAAACATCATTGTCGTTCTGTTACTGTAGTCCCCTCGTTTAGAGGATTGCCA              21060

     L   Y   N   T   D   M   S   F   I   F   S   H   E   V   M   L   L   R   I   Q
     TTATATAATACTGATATGTCTTTTATCTTTAGCCATGAAGTTATGTTATTAAGGATACAA              21120

     N   N   L   A   K   R   S   S   R   F   L   K   R   T   F   D   I   V   C   S
     AATAACTTGGCTAAAAGGTCGTCCCGTTTTCTCAAACGGACATTTGATATTGTTTGTTCA              21180

     I   M   I   L   I   I   A   S   P   L   M   I   Y   L   W   Y   K   V   T   R
     ATAATGATTCTTATAATTGCATCACCACTTATGATTTATCTGTGGTATAAAGTTACTCGA              21240

     D   G   G   P   A   I   Y   G   H   Q   R   V   G   R   H   G   K   L   F   P
     GATGGTGGTCCGGCTATTTATGGTCACCAGCGAGTAGGTCGGCATGGAAAACTTTTTCCA              21300

     C   Y   K   F   R   S   M   V   M   N   S   Q   E   V   L   K   E   L   L   A
     TGCTACAAATTTCGTTCTATGGTTATGAATTCTCAAGAGGTACTAAAAGAACTTTTGGCT              21360

     N   D   P   I   A   R   A   E   W   E   K   D   F   K   L   K   N   D   P   R
     AACGATCCTATTGCCAGGGCTGAATGGGAGAAAGATTTTAAACTGAAAAATGATCCTCGA              21420

     I   T   A   V   G   R   F   I   R   K   T   S   L   D   E   L   P   Q   L   F
     ATCACAGCTGTAGGTCGATTTATACGTAAAACTAGCCTTGATGAGTTGCCACAACTTTTT              21480

     N   V   L   K   G   D   M   S   L   V   G   P   R   P   I   V   S   D   E   L
     AATGTACTAAAAGGTGATATGAGCCTGGTTGGACCACGACCTATCGTTTCGGATGAACTG              21540

     E   R   Y   C   D   D   V   D   Y   Y   L   M   A   K   P   G   M   T   G   L
     GAGCGTTATTGTGATGATGTTGATTATTATTTGATGGCAAAGCCGGGCATGACAGGTCTA              21600

     W   Q   V   S   G   R   N   D   V   D   Y   D   T   R   V   Y   F   D   S   W
     TGGCAAGTGAGTGGGCGTAATGATGTTGATTATGACACTCGTGTTTATTTTGATTCCTGG              21660

     Y   V   K   N   W   T   L   W   N   D   I   A   I   L   F   K   T   A   K   V
     TATGTTAAAAACTGGACGCTTTGGAATGATATTGCCATTCTGTTTAAAACAGCGAAAGTT              21720

                               End of wbaP
     V   L   R   R   D   G   A   Y   *
     GTTTTGCGGCGAGATGGTGCGTAT TAAGCTTACCGAGAAGTACTGAATAATAATTGTATA              21780

     AATTAGCCTGCGTAAAATCTGAACGCATCAATCGCTACCTTAATATCATACCTTTGAGTT              21840
```

Figure 10/16

AACATACTATTCACCTTTAACCTGCCATGACCGTTTGTGGCAGGGTTTCCACACCTGACA   21900

GGAGTATGTAATGTCCAAGCAACAGATCGGCGTCGTCGGTATGGCAGTGATGGGGCGCAA   21960

CCTCGCGCTCAACATCGAAAGCCGTGGTTATACCGTCTCCGTTTTCAACCGCTCCCGTGA   22020

AAAGACCGAAGAAGTGATTGCCGAGAATCCCGGCAAAAAGCTGGTGCCTTATTACACGGT   22080

Figure 10/17